



Katedra
aplikované
matematiky

METODY OPTIMALIZACE

ZDENĚK DOSTÁL,
PETR BEREMLIJSKI

Přepracované vydání

Zdeněk Dostál, Petr Beremlijski
Metody optimalizace

Předmluva

Učební text, který dostáváte do rukou, byl připraven na základě zkušeností z vedení přednášek a cvičení různých variant předmětu Metody optimalizace, který garantuje Katedra aplikované matematiky od svého vzniku pro studenty inženýrských, magisterských i doktorských studijních programů. I když podstatná část studentů přicházela ze studijních programů Fakulty informatiky a elektrotechniky, řada studentů přicházela i z jiných fakult, což se projevilo při výběru témat i metodiky zpracování. Text je proto vhodný pro studenty nejrůznějšího zaměření, a zejména pro ty, kteří hodlají využít metody optimalizace v technické praxi.

Úvod o díle

Naším přáním je uvést čtenáře do základních metod řešení úloh „hladké“ či „téměř hladké“ optimalizace, tj. metod hledání minima funkcí více proměnných, které jsou všude nebo téměř všude (co je zde „téměř“, je specifikováno v textu) spojitě i s některými derivacemi, a to na množině, která je typicky popsána pomocí rovností a nerovností. Optimalizační úlohy vznikají při řešení praktických úloh, když chceme dosáhnout co nejlepšího výsledku v rámci našich možností. Ve výkladu se zaměříme na důležité speciální případy, zejména na konvexní programování, pro které je teorie často poměrně jednoduchá a které jsou důležité pro pochopení řešení praktických problémů. Snažíme se též rozlišovat, co je důležité pro pochopení teorie a co je skutečně důležité vědět pro řešení reálných úloh. Teorii zde chápeme jako nástroj pro výběr metod řešení konkrétních technických problémů, výběr software, ocenění správnosti a spolehlivosti řešení, a také jako nástroj pro studium dalších předmětů. Například význam Karush-Kuhn-Tuckerových podmínek je ilustrován na odvození Snellova zákona. Při psaní textu jsme využili i našich zkušeností z řešení rozsáhlých optimalizačních úloh.

Pro porozumění výkladu je potřebná znalost poznatků z lineární algebry, z diferenciálního počtu reálných funkcí jedné a více reálných proměnných, numerických metod a vhodná je též znalost základů práce s Matlabem. Podle našich zkušeností můžeme před samotným čtením tohoto učebního textu studentům doporučit, aby se seznámili s texty [6], [10], [15] a [9]. Tento text byl inspirován z velké části knihou [5] a dále knihami [1], [12] a [14].

Autoři textu chtějí poděkovat Ondřeji Zjevíkovi za pečlivé přečtení textu a řadu cenných připomínek.

O textu

Nejdůležitější tvrzení jsou graficky zvýrazněna. Věty, definice a algoritmy jsou uvedeny v rámečku. Konce důkazů jsou vyznačeny prázdným čtverečkem (\square), konce řešení příkladů plným trojúhelníčkem (\blacktriangle). Text byl vysázen pomocí sázecího systému $\text{T}_{\text{E}}\text{X}$ ve formátu pdf $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$.

O projektu

Text, který právě čtete, vznikl v rámci řešení projektu „Matematika pro inženýry 21. století – inovace výuky matematiky na technických školách v nových podmínkách rychle se vyvíjející informační a technické společnosti“. Projekt byl řešen na Vysoké škole báňské - Technické univerzitě v Ostravě a Západočeské univerzitě v Plzni v období 2009 – 2012.

Cílem projektu byla inovace matematických a některých odborných kurzů na technických vysokých školách s cílem získat zájem studentů, zvýšit efektivnost výuky, zpřístupnit prakticky aplikovatelné výsledky moderní matematiky a vytvořit předpoklady pro efektivní výuku inženýrských předmětů.

V rámci projektu jsme vytvořili 41 nových výukových materiálů z oblastí matematické analýzy, lineární algebry, numerických metod, metod optimalizace, diskrétní matematiky, teorie grafů, statistiky a několika odborných kurzů. Všechny hotové výukové materiály jsou volně k dispozici na webových stránkách projektu <http://mi21.vsb.cz>.

V Ostravě 1. 4. 2012

Zdeněk Dostál a Petr Beremlijski

Předmluva k přepracovanému vydání

Původní text z roku 2012 byl po 6 letech od svého vzniku v řadě míst doplněn či opraven. Také došlo k úpravám všech obrázků. Věříme, že tyto úpravy pomohou čtenáři k většímu porozumění základních metod řešení úloh „hladké“ či „téměř hladké“ optimalizace. Autoři přepracovaného vydání chtějí poděkovat Marii Sadowské za pečlivé přečtení textu a řadu cenných připomínek a oprav.

V Ostravě 1. 10. 2018

Zdeněk Dostál a Petr Beremlijski

Další drobné opravy byly provedeny na jaře 2023.

Z. Dostál a P. Beremlijski

Obsah

Předmluva	iii
Předmluva k přepracovanému vydání	v
1 Optimalizační úlohy	1
1.1 Úloha optimalizace	1
1.2 Klasifikace optimalizačních úloh	4
2 Podmínky minima pro minimalizaci bez omezení	6
2.1 Taylorova věta	6
2.2 Podmínky minima kvadratické funkce	8
2.3 Podmínky lokálního minima diferencovatelné funkce	9
2.4 Konvexita	10
3 Minimalizace bez derivací	14
3.1 Bisekce	15
3.2 Fibonacciovské dělení a zlatý řez	18
Příklady k procvičení	21
4 Metody spádových směrů	22
4.1 Gradientní metoda pro kvadratickou funkci	22
4.2 Gradientní metoda pro nelineární funkce	25
4.3 Aproximace gradientu pomocí konečných diferencí	27
4.4 Ukončovací podmínky	28
Příklady k procvičení	28
5 Metoda sdružených gradientů	31
5.1 Metoda sdružených směrů pro kvadratické funkce	31
5.2 Metoda sdružených gradientů pro kvadratické funkce	34
5.3 Rychlost konvergence	38
5.4 Metoda sdružených gradientů pro nelineární funkce	39
Příklady k procvičení	40

6	Newtonova metoda	42
6.1	Newtonova metoda	42
6.2	Kvadratická rychlost konvergence	43
6.3	Newtonova metoda s modifikací Hessiánu	44
	Příklady k procvičení	46
7	Trust region	49
7.1	Penalizační lemma	50
7.2	Levenberg-Marquardtova implementace	51
	Příklady k procvičení	52
8	Metody aproximace Hessiánu - BFGS	54
8.1	Heuristické odvození základních myšlenek metody BFGS	54
8.2	Minimalizační lemma	55
8.3	Symetrická pozitivně definitní aproximace Hessiánu	56
8.4	Algoritmus BFGS	59
	Příklady k procvičení	59
9	Podmínky minima pro úlohy s omezením ve tvaru rovností	62
9.1	Geometrická ilustrace podmínky minima	62
9.2	Úloha kvadratického programování s rovnostmi a Lagrangián	63
9.3	Podmínky minima pro kvadratické programování s rovnostmi	65
9.4	Nutné podmínky pro úlohy nelineárního programování	66
9.5	Odvození Snellova zákona lomu	72
10	Metody řešení úloh s rovnostmi	75
10.1	Vnější penalta	75
10.2	Metoda oboru hodnot pro kvadratické programování	77
10.3	Metoda nulového prostoru pro kvadratické programování	78
10.4	Metody Uzawova typu pro kvadratické programování	78
10.5	SMALE	81
	Příklady k procvičení	82
11	Podmínky minima pro úlohy s omezením ve tvaru nerovností	83
11.1	Geometrická ilustrace podmínky minima	83
11.2	Úloha kvadratického programování s nerovnostmi a Lagrangián	84
11.3	Podmínky minima pro kvadratické programování s nerovnostmi	87
11.4	Nutné podmínky pro úlohy nelineárního programování	88
11.5	Klasifikace omezení pro úlohy s omezeními ve tvaru rovností a nerovností	98

12	Metody řešení úloh s nerovnostmi	103
12.1	Metoda vnitřního bodu - bariéry	103
12.2	Metoda aktivních (pracovních) množin - vnější penalta	105
12.3	SMALBE	106
	Příklady k procvičení	106
13	Dualita v konvexním programování	108
13.1	Dualita pro kvadratické programování s rovnostními vazbami	108
13.2	Dualita pro kvadratické programování s nerovnostními vazbami	110
13.3	Vizualizační lemma	113
14	Lineární programování	115
14.1	Formulace úlohy lineárního programování	115
14.2	Podmínky minima a dualita	116
14.3	Metoda vnitřního bodu	117
14.4	Geometrie simplexové metody	118
14.5	Konvexní hranice simplexu	120
14.6	Krok simplexové metody	122
15	Úvod do metod nehladké optimalizace	125
15.1	Základní pojmy	125
15.2	Nepřímé metody	131
15.3	Subgradientní metody	131
15.4	Svazkové metody	133
16	Úvod do metod globální optimalizace	139
16.1	Nelder-Meadova metoda simplexů	139
	Literatura	143
	Rejstřík	145

Kapitola 1

Optimalizační úlohy

V této úvodní kapitole se seznámíme především s formulací optimalizačních úloh, jejichž řešením se v tomto textu budeme zabývat. Ukážeme si některé příklady optimalizačních úloh a pro obecnou orientaci čtenáře uvedeme i přehlednou klasifikaci optimalizačních úloh. Z historických důvodů se některé optimalizační úlohy nazývají také úlohy matematického programování.

1.1 Úloha optimalizace

Optimalizační úlohy vznikají typicky při řešení praktických úloh, když chceme dosáhnout co nejlepšího výsledku v rámci našich možností. Co je nejlepší posuzujeme pomocí *cenové funkce* či *cenového funkcionálu*, což je v obecném případě zobrazení $f : \mathcal{D} \rightarrow \mathbb{R}$ definované na množině \mathcal{D} prvků vhodného vektorového prostoru. Extrémy cenového funkcionálu často hledáme na *přípustné množině* $\Omega \subseteq \mathcal{D}$. Jelikož minimum funkce f je stejné jako maximum funkce $-f$, můžeme za obecnou úlohu optimalizace považovat bez újmy na obecnosti problém najít $\bar{x} \in \Omega$ tak, aby platilo

$$f(\bar{x}) \leq f(x), \quad x \in \Omega. \quad (1.1)$$

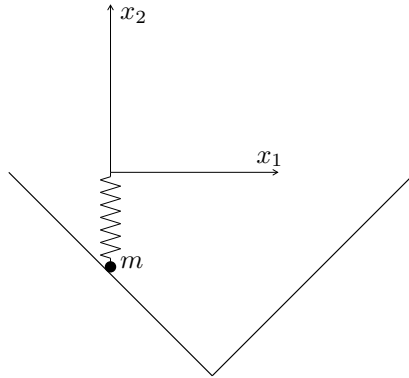
Řešení problému (1.1) nazýváme také *globální řešení*.

Příkladem takové optimalizační úlohy je třeba problém nalezení rovnovážné polohy kuličky o hmotnosti m zavěšené na pružině, která visí nad překážkou. Tento problém můžeme vyjádřit jako minimalizaci funkce potenciální energie tohoto systému. Přípustná množina nám popisuje nepronikání kuličky do překážky. Tuto úlohu pak zapíšeme jako úlohu najít $\bar{x} \in \Omega$ tak, aby platilo

$$f(\bar{x}) \leq f(x), \quad x \in \Omega, \quad (1.2)$$

kde $f(x) = x_1^2 + x_2^2 + mx_2$ a $\Omega = \{x \in \mathbb{R}^2 : x_1 + x_2 + 1 \geq 0, -x_1 + x_2 + 3 \geq 0\}$ (viz obrázek 1.1).

Odkazy na takto obecně formulované úlohy najdeme již ve starověku. Patří mezi ně i úloha, která je pojmenovaná podle princezny Dido z řecké mytologie. Didonina

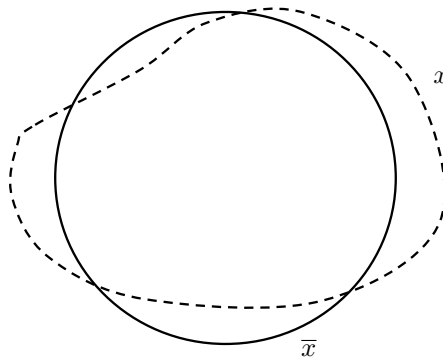


Obr. 1.1: Rovnovážný stav kuličky na pružině

úloha se nazývá také *isoperimetrický problém*. Dido měla dostat pruh kůže z jednoho vola s tím, ať si vybere pozemek, který se dá tímto pruhem omezit. Snaha získat co největší pozemek vede na problém najít spojitou křivku dané délky, která ohraničuje největší plochu. Přírozenou přípustnou množinu Ω zde tvoří množina hladkých uzavřených křivek $x : \langle 0, 1 \rangle \rightarrow \mathbb{R}^2$ dané délky, $x(0) = x(1)$, které tvoří hranici jednoduše souvislé oblasti Ω_x . Cenová funkce pak má tvar

$$f(x) = \int_{\Omega_x} f(s) d\Omega.$$

Pro detaily viz obrázek 1.2.



Obr. 1.2: Didonina úloha

Didonina úloha je velmi specifickou úlohou variačního počtu, kdy neznámou hledáme ve vhodném prostoru funkcí a řešení je možno najít analyticky metodami. Analyticky lze vyřešit i úlohu najít tvar (průběh vertikálního rozměru $x(d)$) nejlehčího vetknutého nosníku obdélníkového průřezu $a \times x(d)$, zatíženého na volném

konci (ve vzdálenosti $d = \ell$ od vetknutí) osamělou silou F . Tuto úlohu zformuloval a částečně vyřešil Galilei ve své knize Dvě nové nauky z roku 1638.

V tomto textu se zaměříme na numerické metody řešení optimalizačních úloh definovaných na \mathbb{R}^n , kdy cenová funkce f je reálná funkce n proměnných, která je spojitá i s některými derivacemi. V tomto případě nazýváme argumenty f také *návrhové proměnné*. Dále budeme předpokládat, že omezení na návrhové proměnné jsou určena *přípustnou množinou* $\Omega \subseteq \mathcal{D}$, kde Ω je definována pomocí hladkých vazebních funkcí $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, m$,

$$\Omega = \{x \in \mathcal{D} : h_{\mathcal{E}}(x) = 0 \quad a \quad h_{\mathcal{I}}(x) \leq 0\}, \quad (1.3)$$

kde \mathcal{I} a \mathcal{E} je rozklad množiny indexů $\{1, \dots, m\}$. Stručný zápis (1.3) lze tedy rozepsat ve tvaru

$$\Omega = \{x \in \mathcal{D} : h_i(x) = 0, \quad i \in \mathcal{E} \quad a \quad h_i(x) \leq 0, \quad i \in \mathcal{I}\}.$$

V případě, že pro Ω platí $\Omega \subseteq \mathcal{D}$ a $\Omega \neq \mathcal{D}$ mluvíme o úlohách *optimalizace s omezením*, příp. o *omezené optimalizaci*. V případě, že pro Ω platí $\Omega = \mathcal{D}$ mluvíme o úlohách *optimalizace bez omezení*, příp. o *neomezené optimalizaci*.

Zaměření na minimalizaci funkcí je užitečné i pro řešení variačních úloh, které nelze vyřešit analyticky, což zahrnuje velkou většinu praktických úloh, jako je úloha najít tvar pólového nástavce elektromagnetu, který maximalizuje homogenitu elektromagnetického pole dané intensity ve vymezené oblasti. Takové úlohy se typicky aproximují úlohou minimalizace funkce více proměnných s velkým počtem neznámých (až desítky a stovky milionů) nebo s velmi náročným vyhodnocením cenové funkce, ke kterému je potřeba najít numerické řešení příslušné stavové rovnice. Efektivní algoritmy minimalizace funkcí více proměnných jsou proto důležité i pro numerické řešení variačních úloh.

Lokální a globální řešení

Úloha najít globální řešení může být příliš ambiciózní, proto má smysl zabývat se hledáním *lokálního řešení*, které minimalizuje pouze zúžení cenové funkce na nějakém okolí \mathcal{U} . Přesněji řečeno, je-li $\bar{x} \in \Omega$ lokální řešení problému (1.1), pak existuje okolí \mathcal{U} bodu \bar{x} tak, aby platilo

$$f(\bar{x}) \leq f(x), \quad x \in \Omega \cap \mathcal{U}.$$

Lokální řešení je užitečné pro zlepšení našeho tipu na řešení.

Obecně je velmi obtížné najít globální řešení, avšak existuje několik případů, kdy to je možné díky speciální struktuře problému. Důležitý případ takové situace nastane, když cenová funkce f i přípustná množina Ω jsou konvexní; v tomto případě mluvíme také o *konvexním programování*.

1.2 Klasifikace optimalizačních úloh

Metody řešení praktických optimalizačních úloh jsou často detailně rozpracovány pro řešení užší třídy problémů se speciální strukturou. My se zde budeme zabývat především řešením úloh *spojité optimalizace*, kdy jak cenová funkce f i funkce popisující rovnosti a nerovnosti h jsou definované na oblasti $\mathcal{D} \subseteq \mathbb{R}^n$. Zvláštní pozornost budeme věnovat *optimalizaci bez omezení*, kdy $\Omega = \mathbb{R}^n$. V rámci spojité optimalizace dále rozlišujeme několik typů problémů, pro které jsou rozpracovány speciální algoritmy numerického řešení.

Lineární programování

Obecná úloha lineárního programování je určena lineární funkcí určenou vektorem $b \in \mathbb{R}^n$ a lineárními omezeními určenými maticí $B \in \mathbb{R}^{m \times n}$ a vektorem $c \in \mathbb{R}^m$. Lze ji zapsat ve tvaru:

$$\text{najdi } \bar{x} \in \Omega \text{ tak, aby } b^T \bar{x} \leq b^T x \text{ pro všechna } x \in \Omega,$$

kde

$$\Omega = \{x \in \mathbb{R}^n : B_{\mathcal{E}}x = c_{\mathcal{E}} \quad \text{a} \quad B_{\mathcal{I}}x \leq c_{\mathcal{I}}\}.$$

První metody řešení úloh lineárního programování byly navrženy před druhou světovou válkou pro účely plánování L. V. Kantorovičem. Řešení speciální úlohy lineárního programování, tzv. dopravní úlohy, sehrálo významnou roli při plánování invaze v Normandii. I když to není na první pohled patrné, některé metody řešení úloh lineárního programování mají diskretní charakter, neboť postupně procházejí konvexní hranici, která typicky obsahuje obrovské množství bodů. Nalezení tzv. elipsoidového algoritmu s polynomiální složitostí N. Karmarkarem v roce 1984 bylo významnou událostí.

Kvadratické programování

Obecná úloha kvadratického programování se od úlohy lineárního programování liší kvadratickou cenovou funkcí, která je určena symetrickou maticí $A \in \mathbb{R}^{n \times n}$ a vektorem $b \in \mathbb{R}^n$. Lze ji zapsat ve tvaru:

$$\text{najdi } \bar{x} \in \Omega, \text{ aby } \frac{1}{2} \bar{x}^T A \bar{x} - b^T \bar{x} \leq \frac{1}{2} x^T A x - b^T x \text{ pro všechna } x \in \Omega,$$

kde Ω je definovaná lineárními rovnostmi a nerovnostmi stejně jako u lineárního programování.

Metody řešení kvadratického programování jsou propojeny s metodami řešení lineárních rovnic. Je-li A pozitivně definitní matice a $\Omega = \mathbb{R}^n$, pak řešení úlohy kvadratického programování je totožné s řešením soustavy lineárních rovnic $Ax = b$. Pro

některé prakticky důležité třídy úloh kvadratického programování existují algoritmy s asymptoticky lineární složitostí, což umožňuje řešit takové úlohy s desítkami až stovkami milionů neznámých.

Nelineární programování

Pokud není cenová funkce ani lineární, ani kvadratická, ale spojitá se spojitou alespoň první derivací, spadá optimalizační problém mezi úlohy nelineárního programování. Zvláštní případ nelineárního programování tvoří úlohy bez omezení, tj. když $\Omega = \mathbb{R}^n$, pro které existují speciální algoritmy. Pokud o cenové funkci ani o vazbách nic nevíme, jedná se však i v tomto případě o obtížnou úlohu. Rozsáhlejší úlohy nelineárního programování lze řešit pouze s použitím algoritmů, které využívají specifické charakteristiky problému.

Nehladká optimalizace

Úlohu nehladké optimalizace dostaneme v případě, že jsou cenová funkce i vazební funkce spojitě a se spojitými alespoň prvními derivacemi až na nějakou „malou“ podmnožinu definičního oboru. V tomto případě se dá tato částečná hladkost využít k odvození efektivních algoritmů. Úlohy nehladké optimalizace jsou obecně obtížnější než úlohy hladké nelineární optimalizace. První obtíž spočívá v komplikovaném popisu podmínek minima, neboť není zaručeno, že funkce definující nehladký optimalizační problém jsou v řešení diferencovatelné.

Globální optimalizace

Jak už jsme se zmínili, zaručené globální minimum nelineární funkce lze najít jen výjimečně. Nejobvyklejší strategií je v tomto případě systematické prohledávání definičního oboru, často s použitím vhodné heuristiky nebo dodatečné informace o cenové funkci a o vazebních funkcích.

Kapitola 2

Podmínky minima pro minimalizaci bez omezení

Dříve než se pustíme do řešení optimalizačních úloh, ukážeme si, jak poznat minimum hladké funkce bez omezení, tedy řešení úlohy

$$\min_{x \in \mathbb{R}^n} f(x). \quad (2.1)$$

Budeme při tom předpokládat, že f je dvakrát spojitě diferencovatelná funkce. V některých důležitých případech si ukážeme, jak lze z míry narušení podmínek minima poznat, jak jsme daleko od minima. Zvláštní pozornost budeme věnovat případům, kdy je zaručena jednoznačnost minima.

Podmínky minima mají význam pro zodpovězení základních teoretických otázek optimalizace. Jejich znalost je užitečná i mimo optimalizaci. Uplatní se na příklad při studiu fyziky a technických aplikací. I když je „ruční“ řešení akademických úloh velmi často založeno na řešení rovnic (a nerovnic) které formálně popisují podmínky minima, počítačové řešení rozsáhlejších praktických úloh je obvykle založeno na jiných postupech, popsanych v dalších kapitolách.

2.1 Taylorova věta

Základní matematický nástroj pro studium řešení optimalizačních problémů je Taylorova věta o aproximaci funkce polynomem v okolí daného bodu. Budeme ji používat v následujícím tvaru.

Věta 2.1. (Taylor) Necht $f : \mathbb{R}^n \rightarrow \mathbb{R}$ je spojitě diferencovatelná a $x, d \in \mathbb{R}^n$. Potom existuje $\alpha \in \mathbb{R}$ tak že

$$f(x + d) = f(x) + \nabla f(x + \alpha d)^T d = f(x) + \nabla f(x)^T d + o(\|d\|), \quad (2.2)$$

kde $\lim_{t \rightarrow 0} o(t)/t = 0$. Je-li navíc f dvakrát spojitě diferencovatelná, potom

$$f(x + d) = f(x) + \nabla f(x)^T d + \frac{1}{2} d^T \nabla^2 f(x)^T d + o(\|d\|^2). \quad (2.3)$$

Taylorova věta nám umožňuje nahradit přírůstek funkce f na okolí daného bodu x lineárním modelem (diferenciálem)

$$\ell(d) = \ell_x(d) = g^T d, \quad g = g(x) = \nabla f(x) = \left[\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right]^T$$

nebo kvadratickým modelem

$$m(d) = m_x(d) = g^T d + \frac{1}{2} d^T H d, \quad H = H(x) = \nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{bmatrix}.$$

Pokud je f kvadratická funkce daná předpisem

$$f(x) = \frac{1}{2} x^T A x - b^T x, \quad A = A^T, \quad (2.4)$$

pak $\nabla f(x) = Ax - b$, $\nabla^2 f(x) = A$, a přírůstek má tvar

$$m_x(d) = f(x + d) - f(x) = (Ax - b)^T d + \frac{1}{2} d^T A d. \quad (2.5)$$

Pomocí Taylorovy věty můžeme poznat, zda je daný vektor d směrem poklesu, tj. zda pro každé dostatečně malé $\epsilon > 0$

$$f(x + \epsilon d) - f(x) < 0.$$

Pomocí (2.2) dostaneme

$$f(x + \epsilon d) - f(x) = \epsilon g^T d + o(\|\epsilon d\|) = \epsilon \left(g^T d + \frac{o(\|\epsilon d\|)}{\epsilon} \right).$$

Jelikož z definice o vyplývá, že $o(\|\epsilon d\|)/\epsilon$ je omezená pro dostatečně malá ϵ , bude výraz v závorce záporný pro všechna dostatečně malá ϵ , pokud $g^T d < 0$. Dostali jsme tak dodatečnou podmínku pro směr poklesu.

2.2 Podmínky minima kvadratické funkce

Následující věta nám dává nutné a postačující algebraické podmínky, které splňují řešení problému

$$\min_{x \in \mathbb{R}^n} f(x), \quad (2.6)$$

kde f je kvadratická funkce definovaná (2.4).

Věta 2.2. *Nechť f je kvadratická funkce (2.4) definovaná symetrickou maticí $A \in \mathbb{R}^{n \times n}$ a $b \in \mathbb{R}^n$. Pak platí následující tvrzení:*

(i) *Vektor \bar{x} je řešení problému (2.6), právě když A je pozitivně semidefinitní a*

$$\nabla f(\bar{x}) = A\bar{x} - b = o. \quad (2.7)$$

(ii) *Problém (2.6) má jediné řešení, právě když A je pozitivně definitní.*

Důkaz. (i) Jestliže \bar{x} a d značí libovolné n -vektory a $\alpha \in \mathbb{R}$, potom můžeme použít Taylorův rozvoj (2.5) a dostaneme

$$f(\bar{x} + \alpha d) - f(\bar{x}) = \alpha(A\bar{x} - b)^T d + \frac{\alpha^2}{2} d^T A d.$$

Předpokládejme nejprve, že \bar{x} je řešení (2.6), takže pravá strana rovnice je nezáporná pro každé α a d . Pro α dostatečně velké a $d \in \mathbb{R}^n$ libovolné, ale pevně zvolené vyplývá z nezápornosti pravé strany $d^T A d \geq 0$, takže A je pozitivně semidefinitní. Na druhé straně, pro α dostatečně malé je znaménko výrazu na pravé straně určeno lineárním členem, takže z nezápornosti pravé strany plyne $(A\bar{x} - b)^T d = 0$ pro každé $d \in \mathbb{R}^n$. Tedy $A\bar{x} - b = o$.

Je-li A pozitivně semidefinitní a \bar{x} splňuje (2.7), potom pro každé $d \in \mathbb{R}^n$ platí

$$f(\bar{x} + d) - f(\bar{x}) = \frac{1}{2} d^T A d \geq 0,$$

tedy \bar{x} je řešení (2.6).

(ii) Je-li \hat{x} jediné řešení minimalizačního problému (2.6), potom podle (i) je A pozitivně semidefinitní a \hat{x} je jediný vektor který splňuje $A\hat{x} = b$. Tedy A je regulární a pozitivně semidefinitní, tj. pozitivně definitní. Na druhé straně, jestliže A je pozitivně definitní, pak je regulární a gradientní podmínku (2.7) splňuje jediný vektor. \square

Jestliže si prohlédneme gradientní podmínku (2.7), dostaneme, že problém (2.6) má řešení tehdy a jen tehdy, když A je pozitivně semidefinitní a

$$b \in \mathcal{H}(A)^1. \quad (2.8)$$

Označíme-li si pomocí R matici, jejíž sloupce tvoří bázi nulového prostoru matice A , můžeme si podmínku řešitelnosti úlohy (2.6) přepsat ve tvaru

$$b^T R = o.$$

Tato podmínka má jednoduchou mechanickou interpretaci: je-li mechanický systém v rovnováze, potom vnější síly musí být kolmé k tuhým pohybům.

Je-li dimenze n problému (2.6) velká, může se ukázat, že snaha najít řešení, které splňuje gradientní podmínku (2.7) přesně, je příliš ambiciózní. Přirozený nápad pak vede k hledání slabší podmínky

$$\|\nabla f(x)\| \leq \varepsilon \quad (2.9)$$

s nějakým malým ε . Jestliže x splňuje nepřesnou podmínku (2.9) s dostatečně malým ε a A je regulární, potom x je blízko jediného řešení \hat{x} , neboť

$$\|x - \hat{x}\| = \|A^{-1}A(x - \hat{x})\| = \|A^{-1}(Ax - b)\| \leq \|A^{-1}\| \|\nabla f(x)\|. \quad (2.10)$$

Typické „řešení“ získané pomocí iteračních řešičů je právě nějaké x které splňuje podmínku (2.9). Pomocí Taylorova rozvoje (2.3) dostaneme

$$\begin{aligned} f(x) - f(\hat{x}) &= f(\hat{x} + (x - \hat{x})) - f(\hat{x}) \\ &= f(\hat{x}) + g(\hat{x})^T(x - \hat{x}) + \frac{1}{2}(x - \hat{x})^T A (x - \hat{x}) - f(\hat{x}) \\ &= \frac{1}{2}(x - \hat{x})^T A (x - \hat{x}). \end{aligned}$$

2.3 Podmínky lokálního minima diferencovatelné funkce

Podmínky lokálního minima pro problém (2.1) získáme snadno analýzou lineárního nebo kvadratického modelu f .

Věta 2.3. *Nechť f je dvakrát diferencovatelná funkce z problému (2.1). Pak platí následující tvrzení:*

(i) *Je-li vektor \bar{x} řešením problému (2.1), pak $g(\bar{x}) = \nabla f(\bar{x}) = o$ a $H(\bar{x}) = \nabla^2 f(\bar{x})$ je pozitivně semidefinitní.*

(ii) *Jestliže vektor $g(\bar{x}) = \nabla f(\bar{x}) = o$ a $H(\bar{x}) = \nabla^2 f(\bar{x})$ je pozitivně definitní, pak je v \bar{x} dosaženo lokálního minima f .*

¹ $\mathcal{H}(A)$ označuje obor hodnot matice A .

Důkaz. (i) Je-li \bar{x} řešení problému (2.1), pak žádný směr $d \in \mathbb{R}^n$ není směrem poklesu funkce. Odtud plyne $d^T g = 0$ pro všechny $d \in \mathbb{R}^n$, tedy i $g^T g = \|g\|^2 = 0$, a podle (2.3) i

$$\frac{\alpha^2}{2} d^T H d + o(\alpha^2 \|d\|^2) = \alpha^2 \left(\frac{1}{2} d^T H d + o(\alpha^2 \|d\|^2) / \alpha^2 \right) \geq 0.$$

Jelikož $\lim_{\alpha \rightarrow 0} o(\alpha^2 \|d\|^2) / \alpha^2 = 0$, dostáváme odsud $d^T H d \geq 0$.

(ii) Je-li H pozitivně definitní a $g(\bar{x}) = \nabla f(\bar{x}) = 0$, pak podle definice $o(\cdot)$ pro dostatečně malé α a všechny $d \in \mathbb{R}^n$ platí

$$\frac{\alpha^2}{2} d^T H d + o(\alpha^2 \|d\|^2) = \alpha^2 \left(\frac{1}{2} d^T H d + o(\alpha^2 \|d\|^2) / \alpha^2 \right) > 0.$$

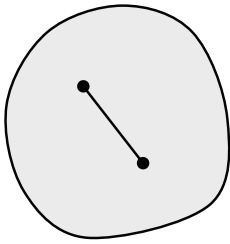
□

Povšimněte si, že věta 2.3 je založena na jednoduchém pozorování: je-li v \bar{x} dosaženo minimum, pak je lineární i kvadratický model přírůstku funkce nezáporný, a je-li lineární i kvadratický model přírůstku funkce kladný na redukovaném okolí \bar{x} , pak je i skutečný přírůstek funkce $\Delta_x(d) = f(\bar{x} + d) - f(\bar{x})$ na redukovaném okolí \bar{x} kladný.

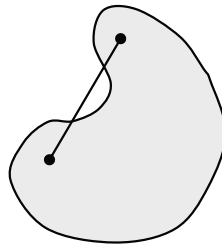
2.4 Konvexita

Mnoho silných tvrzení je možno dokázat pro problémy, které splňují podmínky konvexity. Intuitivně, konvexita je vlastnost množin které obsahují s každou dvojicí bodů také celou úsečku, která je spojuje tak jako na obrázku 2.1. Obdobně množina není konvexní, pokud existuje dvojice bodů z dané množiny, která neobsahuje úsečku, která je spojuje (tak jako na obrázku 2.2).

Definice 2.4. Formálněji, podmnožina Ω množiny \mathbb{R}^n je *konvexní*, jestliže pro každé x a y v Ω a $\alpha \in (0, 1)$ leží vektor $s = \alpha x + (1 - \alpha)y$ také v Ω .



Obr. 2.1: Konvexní množina



Obr. 2.2: Nekonvexní množina

Definice 2.5. Necht x_1, \dots, x_k jsou vektory z \mathbb{R}^n . Jestliže $\alpha_1, \dots, \alpha_k$ jsou takové skaláry, že platí

$$\alpha_i \geq 0, \quad i = 1, \dots, k, \quad \sum_{i=1}^k \alpha_i = 1,$$

potom se vektor $v = \sum_{i=1}^k \alpha_i x_i$ nazývá *konvexní kombinace* vektorů x_1, \dots, x_k .

Definice 2.6. *Konvexní obal* vektorů x_1, \dots, x_k , značený $\text{conv}\{x_1, \dots, x_k\}$, je množina všech konvexních kombinací x_1, \dots, x_k .

Konvexní obal vektorů x_1, \dots, x_k je nejmenší konvexní množina, do které patří x_1, \dots, x_k .

Definice 2.7. *Konvexní hranice* konvexní množiny Ω je množina vektorů z Ω , které nemohou být vyjádřeny jako konvexní kombinace jiných vektorů z Ω .

Konvexní hranice uzavřeného čtverce je tvořena jeho čtyřmi rohy, zatímco konvexní hranice uzavřeného kruhu je tvořena celou jeho hranicí. Průnik dvou a více konvexních množin je opět konvexní. Jako příklady konvexních množin si uveďme kruh, čtverec a množiny definované konečnou množinou lineárních rovnic jako $b^T x = c$ nebo nerovnic jako $b^T x \leq c$.

Definice 2.8. Zobrazení $h : \Omega \rightarrow \mathbb{R}$ definované na konvexní množině $\Omega \in \mathbb{R}^n$ se nazývá *konvexní funkce*, je-li její nadgraf konvexní množina, tj.

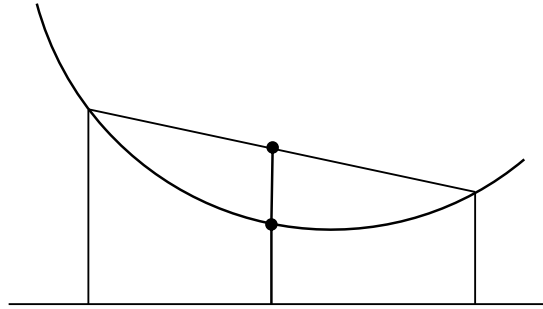
$$h(\alpha x + (1 - \alpha)y) \leq \alpha h(x) + (1 - \alpha)h(y)$$

pro všechna $x, y \in \Omega$ a $\alpha \in (0, 1)$, a *ryze konvexní*, když

$$h(\alpha x + (1 - \alpha)y) < \alpha h(x) + (1 - \alpha)h(y)$$

pro všechna $x, y \in \Omega$, $x \neq y$, a $\alpha \in (0, 1)$.

Pojem konvexní funkce je ilustrován na obr. 2.3.



Obr. 2.3: Konvexní funkce

Uvedme si ještě charakteristiku konvexní kvadratické funkce.

Věta 2.9. *Kvadratická funkce f definovaná v (2.4) s pozitivně semidefinitním Hesiánem A je konvexní. Je-li A pozitivně definitní, je f ryze konvexní.*

Důkaz. Necht $x, y \in \mathbb{R}^n$, $\alpha \in (0, 1)$, a $s = \alpha x + (1 - \alpha)y$. Potom z Taylorova rozvoje (2.3) funkce f v s dostaneme

$$\begin{aligned} f(s) + \nabla f(s)^T(x - s) + \frac{1}{2}(x - s)^T A(x - s) &= f(x), \\ f(s) + \nabla f(s)^T(y - s) + \frac{1}{2}(y - s)^T A(y - s) &= f(y). \end{aligned}$$

Vynásobíme-li první rovnici α , druhou rovnici $1 - \alpha$, a sečteme výsledné rovnice, dostaneme

$$\begin{aligned} f(s) + \frac{\alpha}{2}(x - s)^T A(x - s) + \frac{1 - \alpha}{2}(y - s)^T A(y - s) \\ = \alpha f(x) + (1 - \alpha)f(y). \end{aligned} \tag{2.11}$$

Odtud plyne, že je-li A pozitivně semidefinitní, potom f je konvexní. Jelikož $x = y$ je ekvivalentní $x = s$ a $y = s$, dostáváme pro A pozitivně definitní, že f je ryze konvexní. \square

Konvexní funkce mají důležitou vlastnost, a sice, že každé jejich lokální minimum je globální minimum.

Věta 2.10. *Necht f je konvexní funkce a $\Omega \subseteq \mathbb{R}^n$ konvexní množina. Potom je každé lokální minimum globálním minimem. Je-li navíc f ryze konvexní, pak je dosaženo nejvýše v jediném bodě.*

Důkaz. Necht $\bar{x} \in \Omega$ a $\bar{y} \in \Omega$ jsou lokální minima f vzhledem k Ω a navíc platí $f(\bar{x}) < f(\bar{y})$. Označíme-li si $y_\alpha = \alpha\bar{x} + (1 - \alpha)\bar{y}$ a využijeme toho, že f je konvexní, dostaneme

$$f(y_\alpha) = f(\alpha\bar{x} + (1 - \alpha)\bar{y}) \leq \alpha f(\bar{x}) + (1 - \alpha)f(\bar{y}) < f(\bar{y})$$

pro každé $\alpha \in (0, 1)$. Jelikož platí

$$\|\bar{y} - y_\alpha\| = \alpha\|\bar{y} - \bar{x}\|,$$

pak výše uvedená nerovnost je ve sporu s předpokladem že v \bar{y} je dosaženo lokálního minima.

Zbývající část důkazu tvrzení přenecháváme čtenáři, který má rád analýzu. \square

Poznamenejme, že předcházející tvrzení nic neříká o existenci globálního minima funkce f . Pokud by nás zajímala informace o existenci globálního minima, můžeme použít například následující tvrzení, které uvádíme bez důkazu.

Věta 2.11. *Spojité funkce f na kompaktní množině $\Omega \subseteq \mathbb{R}^n$ má na této množině globální minimum.*

Kapitola 3

Minimalizace bez derivací

V této části se budeme zabývat metodami minimalizujícími funkce bez omezení. Metodami bez derivací (derivative-free methods) máme na mysli metody, které nepoužívají derivace (či gradienty) minimalizovaných funkcí ani aproximace těchto derivací (či gradientů). Tyto problémy můžeme rozdělit na úlohy jednorozměrné optimalizace a úlohy vícerozměrné optimalizace. V případě úloh jednorozměrné optimalizace se v této kapitole budeme zabývat metodami bisekce, Fibonacciovským dělením a metodou zlatého řezu. Všechny tyto metody jsou využitelné nejen pro řešení úloh jednorozměrné optimalizace, ale také pro metody vícerozměrné optimalizace používající strategii tzv. linesearch. V případě úloh vícerozměrné optimalizace se metodám bez derivací budeme věnovat v kapitole 16.

Metody bez použití derivací nebo také metody přímého vyhledávání (direct search methods)¹ byly rozvíjeny a používány zejména v šedesátých letech dvacátého století. Důvodů bylo více. Jedním z nich byla chybějící globalizační strategie pro určování délky kroku pro metody prvního a druhého řádu. Z toho důvodu byly metody nultého řádu na začátku šedesátých let „spolehlivější“ než metody prvního a druhého řádu. Armijo-Goldsteinovy podmínky, které se nyní používají pro zajištění konvergence, byly publikovány až v roce 1966. Dalším důvodem, který platí stále, je to, že metody prvního a druhého řádu nefungují pro všechny optimalizační problémy. Tyto metody selhávají mimo jiné při minimalizaci nediferencovatelných či nespojitých funkcí. Zatímco metody nultého řádu jsou schopny najít minima i těchto funkcí. Třetím důvodem, a také stále přetrvávajícím, je velmi snadná implementace a technická realizace těchto metod. Velká výhoda je v tom, že není třeba hledat ani Hessián ani gradient během iteračního procesu. Z těchto důvodů jsou tyto metody stále v určité míře používány, i přes existenci daleko pokročilejších optimalizačních algoritmů. Podrobný přehled metod z této kapitoly může čtenář nalézt v [11].

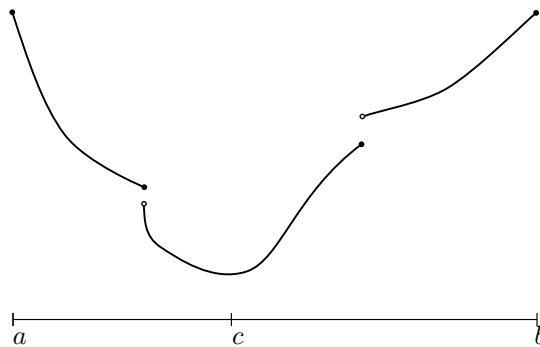
¹Někdy se pro ně používá také označení metody nultého řádu.

3.1 Bisekce

V této i následující podkapitole se budeme zabývat metodami pro hledání minima tzv. *unimodálních* funkcí.

Definice 3.1. O funkci $f: \mathbb{R} \rightarrow \mathbb{R}$ řekneme, že je unimodální v intervalu $\langle a, b \rangle$ ($-\infty \leq a < b \leq \infty$), jestliže existuje $c \in \langle a, b \rangle$ takové, že funkce f je klesající v intervalu $\langle a, c \rangle$ a rostoucí v intervalu $\langle c, b \rangle$.

Poznámka 3.2. Pozorný čtenář si v předcházející definici určitě všiml, že funkce f nemusí být v intervalu $\langle a, b \rangle$ ani diferencovatelná a dokonce ani spojitá (viz obrázek 3.1).



Obr. 3.1: Příklad unimodální funkce

Poznámka 3.3. Je snadné ukázat, že funkce, která je v intervalu $\langle a, b \rangle$ konvexní, je v něm i unimodální.

Věta 3.4. Necht funkce f je unimodální v $\langle a, b \rangle$ a nabývá svého minima v $x^* \in \langle a, b \rangle$. Pak pro libovolné x_1, x_2 , kde $a \leq x_1 < x_2 \leq b$, platí následující tvrzení:

- i) Je-li $f(x_1) < f(x_2)$, pak $x^* < x_2$.
- ii) Je-li $f(x_1) > f(x_2)$, pak $x^* > x_1$.
- iii) Je-li $f(x_1) = f(x_2)$, pak $x_1 < x^* < x_2$.

Důkaz. Důkaz provedeme sporem.

Ve všech případech předpokládejme, že f je unimodální v $\langle a, b \rangle$ a nabývá svého minima v $x^* \in \langle a, b \rangle$. Uvažujme dvojici x_1, x_2 splňující podmínky věty 3.4.

- i) Dále platí $f(x_1) < f(x_2) \wedge x^* \geq x_2$. Odtud plyne, že f je klesající v $\langle a, x^* \rangle$ a odtud $f(x_2) < f(x_1)$, což je spor s předpokladem.

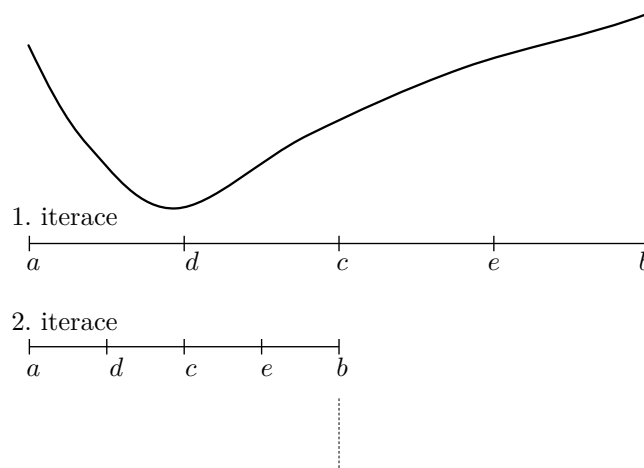
- ii) Dále platí $f(x_1) > f(x_2) \wedge x^* \leq x_1$. Odtud plyne, že f je rostoucí v $\langle x^*, b \rangle$ a odtud $f(x_2) > f(x_1)$, což je spor s předpokladem.
- iii) Dále platí $f(x_1) = f(x_2) \wedge (x^* \leq x_1 \vee x^* \geq x_2)$. Odtud plyne, že buď (v případě, že $x^* \leq x_1$) je f rostoucí v $\langle x^*, b \rangle$ a odtud $f(x_2) > f(x_1)$, což je spor s předpokladem. Anebo (v případě, že $x^* \geq x_2$) je f klesající v $\langle a, x^* \rangle$ a odtud $f(x_2) < f(x_1)$, což je opět spor s předpokladem.

□

Metoda bisekce je určena pro hledání minima funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ v intervalu $\langle a, b \rangle$. Předpokládejme, že tato funkce je v tomto intervalu unimodální. Tzn. že má v tomto intervalu právě jedno minimum. S využitím věty 3.4 lze snadno navrhnout algoritmus pro hledání minima unimodální funkce. Tento algoritmus bude generovat posloupnost intervalů

$$I_0 \supset I_1 \supset I_2 \supset \dots \supset I_{k-1} \supset I_k \supset \dots, \quad I_k = \langle a_k, b_k \rangle.$$

Všechny intervaly dané posloupnosti obsahují hledané minimum funkce f . V případě metody bisekce bude poměr délek dvou po sobě jdoucích intervalů 1 ku 2. Zvolme v intervalu $I_{k-1} = \langle a_{k-1}, b_{k-1} \rangle$ následující body $c_{k-1} := \frac{a_{k-1} + b_{k-1}}{2}$, $d_{k-1} := \frac{a_{k-1} + c_{k-1}}{2}$ a $e_{k-1} := \frac{c_{k-1} + b_{k-1}}{2}$. Pokud porovnáme funkční hodnoty v bodech c_{k-1} , d_{k-1} , e_{k-1} a využijeme tvrzení věty 3.4, získáme interval I_k (viz obrázek 3.2). Algoritmus ukončíme v okamžiku, kdy velikost intervalu I_k bude menší než požadovaná hodnota. Minimum funkce f bude aproximováno středem tohoto intervalu c_k . Podrobný algoritmus je uveden níže.



Obr. 3.2: Algoritmus bisekce

Algoritmus 3.1 Bisekce.

Krok 0. {Inicializace.}
 $\varepsilon > 0$ (přesnost), $a_0, b_0, c_0 := \frac{a_0+b_0}{2}, k := 0$

Krok 1. {Cyklus bisekce.}
while $b_k - a_k \geq \varepsilon$
 $d_k := \frac{a_k+c_k}{2}$
 $e_k := \frac{c_k+b_k}{2}$
if $f(d_k) < f(c_k)$
 $a_{k+1} := a_k$
 $b_{k+1} := c_k$
 $c_{k+1} := d_k$
elseif $f(c_k) < f(e_k)$
 $a_{k+1} := d_k$
 $b_{k+1} := e_k$
 $c_{k+1} := c_k$
else
 $a_{k+1} := c_k$
 $b_{k+1} := b_k$
 $c_{k+1} := e_k$
end
 $k := k + 1$
end while

Krok 2. {Dosad (přibližné) řešení.}
 $x_k := c_k$

Příklad 3.5. Provedte dvě iterace metody bisekce pro minimalizaci funkce $f(x) = x^2 - 0,5x + 0,0625$ na intervalu $\langle 0, 1 \rangle$.

Řešení. $k = 0$:

$a_0 = 0, b_0 = 1, c_0 := \frac{a_0+b_0}{2} = 0,5, d_0 := \frac{a_0+c_0}{2} = 0,25, e_0 := \frac{c_0+b_0}{2} = 0,75.$
 $f(c_0) = 0,0625, f(d_0) = 0, f(e_0) = 0,25.$

$k = 1$:

$a_1 := a_0 = 0, b_1 := c_0 = 0,5, c_1 := d_0 = 0,25, d_1 := \frac{a_1+c_1}{2} = 0,125, e_1 := \frac{c_1+b_1}{2} = 0,375.$
 $f(c_1) = f(d_0) = 0, f(d_1) = 0,015625, f(e_1) = 0,015625.$

$k = 2$:

$a_2 := d_1 = 0,125, b_2 := e_1 = 0,375, c_2 := c_1 = 0,25.$

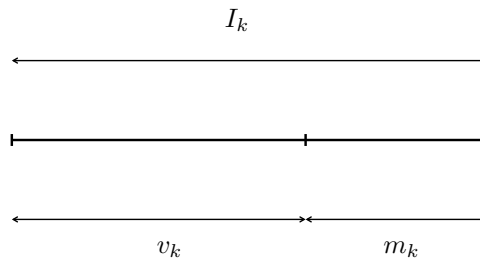
Po dvou iteracích tedy víme, že minimum leží v intervalu $\langle 0,125, 0,375 \rangle$. $x_2 := c_2 = 0,25$ aproximuje minimum funkce $f(x)$ ▲

3.2 Fibonacciovské dělení a zlatý řez

Nejprve připomeňme, že *Fibonacciova posloupnost* je posloupnost zadaná rekurentním předpisem

$$F_{n+2} = F_{n+1} + F_n, \quad F_0 = F_1 = 1.$$

Tedy prvních několik členů této posloupnosti má hodnotu: $F_0 = F_1 = 1$, $F_2 = 2$, $F_3 = 3$, $F_4 = 5$, $F_5 = 8$, $F_6 = 13$, $F_7 = 21$, $F_8 = 34$, $F_9 = 55$, $F_{10} = 89$, \dots Nyní si ukážeme, jak pomocí této posloupnosti sestrojít sekvenci intervalů, ve kterých leží minimum unimodální funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ v intervalu $\langle 0, 1 \rangle$. Neprve zvolme počet členů této sekvence jako $N \in \mathbb{N}$, $N > 2$, tím také určíme apriori přesnost aproximace hledaného minima. Necht' $x_1 = \frac{F_{N-1}}{F_N}$ a $x_2 = \frac{F_{N-2}}{F_N}$. Pokud porovnáme funkční hodnoty v bodech x_1 , x_2 a využijeme tvrzení věty 3.4, získáme buď interval $\langle 0, x_1 \rangle$ nebo $\langle x_2, 1 \rangle$. V tomto intervalu již leží bod x_2 nebo x_1 a dopočteme v něm bod $x_3 = 0 + \frac{F_{N-3}}{F_N}$ nebo $x_3 = 1 - \frac{F_{N-3}}{F_N}$. A takto postupujeme dále až do bodu x_{N-1} . Nevýhodou tohoto postupu je, že pokud se rozhodneme zpřesnit aproximaci hledaného řešení (čehož docílíme zvětšením hodnoty N) musíme všechny body x_1, x_2, \dots, x_{N-1} a odpovídající funkční hodnoty vyčíslit znovu. Tuto nevýhodu můžeme odstranit, pokud zafixujeme zkracování intervalu na jedné hodnotě a to $\bar{\gamma} = \lim \frac{F_{N-1}}{F_N}$. Tuto novou metodu nazveme *metodou zlatého řezu*. Metoda zlatého řezu je založena na rozdělení každého intervalu I_k na dva intervaly v_k a m_k (viz obrázek 3.3) tak, aby



Obr. 3.3: Algoritmus zlatého řezu - nástin dělení intervalu

poměr větší části v_k k menší části m_k byl stejný jako poměr celého intervalu I_k k větší části v_k , tj. platí

$$\frac{v_k}{m_k} = \frac{I_k}{v_k} = \gamma = \text{konst.}, \quad (3.1)$$

$$I_k = v_k + m_k. \quad (3.2)$$

Rovnici (3.2) upravíme na

$$\frac{I_k}{v_k} = 1 + \frac{m_k}{v_k} \quad (3.3)$$

a s využitím (3.1) dostaneme

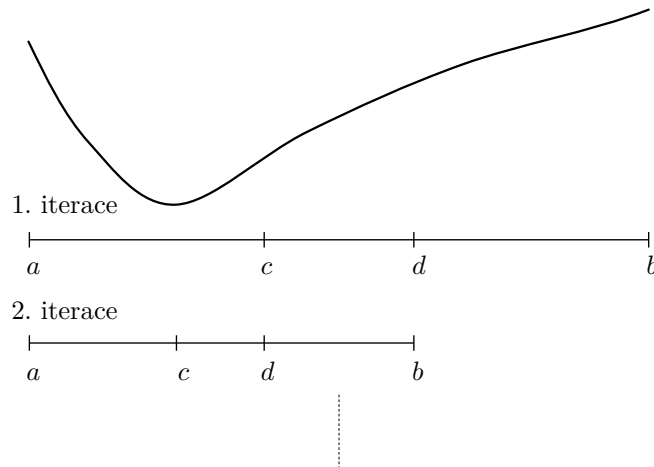
$$\gamma = 1 + \frac{1}{\gamma} \quad (3.4)$$

Z toho snadno odvodíme, že $\gamma = \frac{1+\sqrt{5}}{2}$.

Metoda zlatého řezu je obdobně jako metoda bisekce určena pro hledání minima unimodální funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ v intervalu $\langle a, b \rangle$. Tato metoda opět generuje posloupnost intervalů

$$I_0 \supset I_1 \supset I_2 \supset \dots \supset I_{k-1} \supset I_k \supset \dots, \quad I_k = \langle a_k, b_k \rangle.$$

Všechny intervaly dané posloupnosti obsahují hledané minimum funkce f . Zvolme v intervalu $I_{k-1} = \langle a_{k-1}, b_{k-1} \rangle$ následující body $d_{k-1} := \frac{b_{k-1}-a_{k-1}}{\gamma} + a_{k-1}$ a $c_{k-1} := a_{k-1} + b_{k-1} - d_{k-1}$. Pokud porovnáme funkční hodnoty v bodech c_{k-1} , d_{k-1} a využijeme tvrzení věty 3.4, získáme interval I_k (viz obrázek 3.4). Algoritmus ukončíme v okamžiku, kdy velikost intervalu I_k bude menší než požadovaná hodnota. Minimum funkce f bude aproximováno středem tohoto intervalu. Podrobný algoritmus je uveden níže.



Obr. 3.4: Algoritmus zlatého řezu

Algoritmus 3.2 **Zlatý řez.**

Krok 0. {Inicializace.}

$$\varepsilon > 0 \text{ (přesnost)}, \gamma := \frac{1+\sqrt{5}}{2}, a_0, b_0, d_0 := \frac{b_0-a_0}{\gamma} + a_0, c_0 := a_0 + b_0 - d_0, k := 0$$

Krok 1. {Cyklus zlatého řezu.}

while $b_k - a_k \geq \varepsilon$

if $f(c_k) < f(d_k)$

$$a_{k+1} := a_k$$

$$b_{k+1} := d_k$$

$$d_{k+1} := c_k$$

$$c_{k+1} := a_{k+1} + b_{k+1} - d_{k+1}$$

else

$$a_{k+1} := c_k$$

$$b_{k+1} := b_k$$

$$c_{k+1} := d_k$$

$$d_{k+1} := a_{k+1} + b_{k+1} - c_{k+1}$$

end

$$k := k + 1$$

end while

Krok 2. {Dosad (přibližné) řešení.}

$$x_{k+1} := \frac{a_{k+1} + b_{k+1}}{2}$$

Příklad 3.6. Provedte dvě iterace metody zlatého řezu pro minimalizaci funkce $f(x) = x^2 - 0,5x + 0,0625$ na intervalu $\langle 0, 1 \rangle$.

Řešení. $k = 0$:

$$a_0 = 0, b_0 = 1, d_0 := \frac{b_0-a_0}{\gamma} + a_0 \doteq 0,618, c_0 := a_0 + b_0 - d_0 \doteq 0,382.$$

$$f(c_0) = 0,017424, f(d_0) = 0,135424.$$

$k = 1$:

$$a_1 := a_0 = 0, b_1 := d_0 = 0,618, d_1 := c_0 = 0,382, c_1 := a_1 + b_1 - d_1 \doteq 0,236.$$

$$f(c_1) = 0,000196, f(d_1) = f(c_0) = 0,017424.$$

$k = 2$:

$$a_2 := a_1 = 0, b_2 := d_1 = 0,382, d_2 := c_1 = 0,236, c_2 := a_2 + b_2 - d_2 \doteq 0,146.$$

Po dvou iteracích tedy víme, že minimum leží v intervalu $\langle 0, 0,382 \rangle$. $x_2 := \frac{a_2+b_2}{2} = 0,191$ aproximuje minimum funkce $f(x)$ ▲

Poznámka 3.7. Pokud porovnáme zkracování intervalů v jednotlivých iteracích v případě metody bisekce a zlatého řezu, zjistíme, že v případě bisekce dochází k většímu zkrácení než v případě metody zlatého řezu. Na druhé straně metoda

bisekce vyžaduje v každé iteraci 2 vyčíslení cenové funkce, zatímco metoda zlatého řezu potřebuje v každé iteraci pouze 1 vyčíslení. Pokud porovnáme zkrácení intervalu na 1 vyčíslení cenové funkce, snadno zjistíme, že je metoda zlatého řezu efektivnější.

Příklady k procvičení

1. Nakreslete graf funkce

$$f(x) = x^4 + 8x^3 + 24x^2 + 32x + 16.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}} f(x).$$

Tuto úlohu řešte metodou bisekce a metodou zlatého řezu. Porovnejte rychlost konvergence (tj. počty iterací) a počty vyčíslení minimalizované funkce při použití jednotlivých metod.

2. Nakreslete graf funkce

$$f(x) = \sin^2 x.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}} f(x).$$

Tuto úlohu řešte metodou bisekce a metodou zlatého řezu. Porovnejte rychlost konvergence (tj. počty iterací) a počty vyčíslení minimalizované funkce při použití jednotlivých metod.

3. Nakreslete graf funkce

$$f(x) = \cos^2 x.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}} f(x).$$

Tuto úlohu řešte metodou bisekce a metodou zlatého řezu. Porovnejte rychlost konvergence (tj. počty iterací) a počty vyčíslení minimalizované funkce při použití jednotlivých metod.

Kapitola 4

Metody spádových směrů

Nejdříve se seznámíme s variantami iteračních algoritmů pro řešení úlohy

$$\min_{x \in \mathbb{R}^n} f(x), \quad (4.1)$$

které využívají v každém kroku směr poklesu hodnoty cenové funkce f k nalezení další iterace s menší hodnotou f . Zaměříme se zejména na varianty *gradientní metody*, která využívá jako směr poklesu záporný gradient. Gradientní metoda se nazývá také *metoda největšího spádu*. Alternativní název pochází z pozorování, že minimum diferenciálu hladké funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}$ pro $d \in \mathbb{R}^n$, $\|d\| = 1$ je dosaženo pro $d = -\|\nabla f(x)\|^{-1} \nabla f(x)$, neboť

$$\min_{\|d\|=1} \nabla f(x)^T d \geq - \min_{\|d\|=1} \|\nabla f(x)\| \|d\| = -\|\nabla f(x)\| = \nabla f(x)^T (-\|\nabla f(x)\|^{-1} \nabla f(x)).$$

Jelikož spádové algoritmy jsou velmi jednoduché na implementaci, mají velmi nízkou cenu iterace a jsou robustní, dají se aplikovat na řešení mnoha úloh. Cenou za to je často velký počet iterací, což není překvapivé, neboť typicky využívají pouze lineární model funkce a nepoužívají explicitně informaci z předchozích iterací.

4.1 Gradientní metoda pro kvadratickou funkci

Základy gradientních metod si vyložíme na minimalizaci konvexní kvadratické funkce

$$f(x) = \frac{1}{2} x^T A x - b^T x, \quad (4.2)$$

která je definována symetrickou pozitivně definitní maticí A a vektorem $b \in \mathbb{R}^n$. Použijeme variantu s „optimální“ délkou kroku α_{opt} , tj. s délkou kroku, která minimalizuje hodnotu cenové funkce f v daném směru poklesu d . Jelikož

$$\frac{df(x - \alpha d)}{d\alpha} = \alpha d^T A d + b^T d - x^T A d = \alpha d^T A d - d^T (A x - b) = \alpha d^T A d - d^T g,$$

dostaneme z podmínky minima

$$\alpha_{\text{opt}} d^T A d - d^T g = 0, \quad \alpha_{\text{opt}} = d^T g / d^T A d.$$

Schéma metody největšího spádu pro kvadratickou funkci s délkou kroku minimalizující hodnotu cenové funkce (4.2) vypadá následovně.

Algoritmus 4.1 Gradientní metoda s „optimální“ délkou kroku pro kvadratickou funkci.

Je dána symetrická pozitivně definitní matice $A \in \mathbb{R}^{n \times n}$ a $b \in \mathbb{R}^n$.

Krok 0. {Inicializace.}

Zvolme $\varepsilon > 0$ (přesnost), $x^0 \in \mathbb{R}^n$, polož $g^0 = Ax^0 - b$, $k = 0$

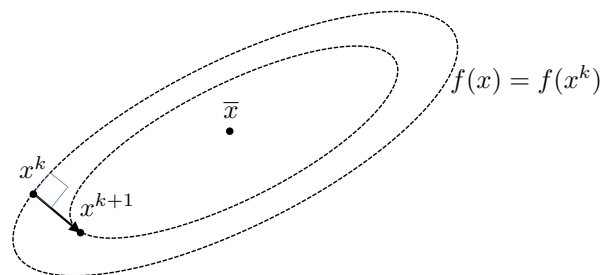
Krok 1. {Cyklus gradientní metody.}

while $\|g^k\| \geq \varepsilon$
 $\alpha_k = (g^k)^T g^k / (g^k)^T A g^k$
 $x^{k+1} = x^k - \alpha_k g^k$
 $g^{k+1} = g^k - \alpha_k A g^k$
 $k = k + 1$
end while

Krok 2. {Dosad (přibližné) řešení.}

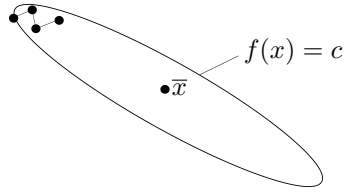
$\tilde{x} = x^k$

Postup gradientní metody s délkou kroku, která maximalizuje pokles cenové funkce, je znázorněn na obr. 4.1.



Obr. 4.1: Gradientní metoda s „optimální“ délkou kroku

Z obr 4.1 je patrna také hlavní příčina pomalé konvergence algoritmu. Je zřejmé, že největší pokles hodnoty cenové funkce ve směru největšího lokálního spádu nezaručuje rozumný globální pokles. Tento aspekt je ještě lépe znázorněn na obr 4.2.



Obr. 4.2: Pomalá konvergence metody největšího spádu

Délka kroku α_{opt} v daném směru d minimalizuje $f(x^k - \alpha d)$ pro $\alpha > 0$, tudíž pro řešení \bar{x} a $\alpha > 0$ platí

$$f(x^k - \alpha_{\text{opt}}d) - f(\bar{x}) \leq f(x^k - \alpha d) - f(\bar{x}).$$

Označíme-li si pro libovolné $x \in \mathbb{R}$ energetickou normu x pomocí

$$\|x\|_A^2 = x^T A x,$$

potom

$$\|x - \bar{x}\|_A^2 = (x - \bar{x})^T A (x - \bar{x}) = (x - A^{-1}b)^T A (x - A^{-1}b) = 2(f(x) - f(\bar{x}))$$

a pro libovolné $\alpha > 0$

$$\|x^{k+1} - \bar{x}\|_A^2 \leq \|x^k - \alpha g^k - \bar{x}\|_A^2.$$

Snadno je vidět, že α_{opt} neminimalizuje vzdálenost od řešení.

Gradientní metodu lze použít i s vhodnou pevnou délkou kroku $\bar{\alpha}$, tedy s volbou $\alpha^k = \bar{\alpha}$, takže

$$x^{k+1} = x^k - \bar{\alpha}g^k.$$

Pomocí $A\bar{x} = b$ a vlastností normy

$$\begin{aligned} \|x^{k+1} - \bar{x}\| &= \|x^k - \bar{\alpha}g^k - \bar{x}\| = \|x^k - \bar{\alpha}(Ax^k - b) - \bar{x} + \bar{\alpha}(A\bar{x} - b)\| \\ &= \|(I - \bar{\alpha}A)(x^k - \bar{x})\| \leq \|I - \bar{\alpha}A\| \|x^k - \bar{x}\| \\ &\leq \max_{\lambda_i \in \sigma(A)} |1 - \bar{\alpha}\lambda_i| \|x^k - \bar{x}\| = \max\{1 - \bar{\alpha}\lambda_{\min}, \bar{\alpha}\lambda_{\max} - 1\} \|x^k - \bar{x}\|, \end{aligned}$$

kde $\sigma(A)$ značí spektrum matice A a λ_{\min} , λ_{\max} jsou extrémní vlastní čísla matice A .

Snadno se ukáže, že absolutní hodnota v posledním výrazu je menší než jedna pro

$$\bar{\alpha} \in (0, 2\lambda_{\max}^{-1})$$

a nabývá nejmenší hodnoty pro

$$1 - \bar{\alpha}\lambda_{\min} = \bar{\alpha}\lambda_{\min} - 1,$$

takže délka kroku $\bar{\alpha}_{\text{opt}}$ je dána předpisem

$$\bar{\alpha}_{\text{opt}} = 2/(\lambda_{\min} + \lambda_{\max}).$$

Pro tuto hodnotu dostaneme lineární rychlost konvergence gradientní metody s optimálním pevným krokem $\bar{\alpha}_{\text{opt}}$

$$\|x^{k+1} - \bar{x}\| \leq \left(1 - \frac{2\lambda_{\min}}{\lambda_{\min} + \lambda_{\max}}\right) \|x^k - \bar{x}\| = \frac{\kappa(A) - 1}{\kappa(A) + 1} \|x^k - \bar{x}\|, \quad (4.3)$$

kde $\kappa(A) = \lambda_{\max}/\lambda_{\min}$ je číslo podmíněnosti matice A .

4.2 Gradientní metoda pro nelineární funkce

Metodu nejmenšího spádu můžeme použít i pro obecnější diferencovatelné funkce, pokud dokážeme najít nebo alespoň aproximovat gradient cenové funkce. Většinou nebudeme schopni najít přesné řešení jednorozměrné minimalizační úlohy. Gradientní metodu budeme proto chápat jako obecné schéma, které můžeme specifikovat i s případným využitím znalostí problému. Obecné schéma metody největšího spádu bez specifikace délky kroku vypadá následovně.

Algoritmus 4.2 Metoda největšího spádu pro nelineární funkce.

Je dána diferencovatelná funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

Krok 0. {Inicializace.}

Zvolme $\varepsilon > 0$ (přesnost), $x^0 \in \mathbb{R}^n$, polož $g^0 = \nabla f(x^0)$, $k = 0$

Krok 1. {Cyklus gradientní metody.}

while $\|g^k\| \geq \varepsilon$

Vyber $\alpha_k > 0$

$x^{k+1} = x^k - \alpha_k g^k$

$g^{k+1} = \nabla f(x^{k+1})$

$k = k + 1$

end while

Krok 2. {Dosad (přibližné) řešení.}

$\tilde{x} = x^k$

Délka kroku

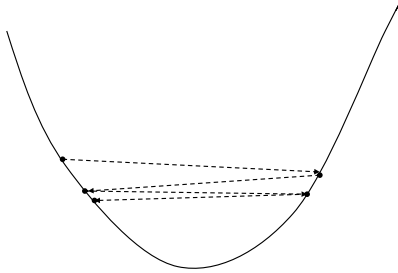
K úplnosti algoritmu je třeba specifikovat délku kroku. Nejpřirozenější volba délky kroku vyplývá z úvahy, že se vyplatí maximálně využít daný směr poklesu obdobně jako jsme to udělali při minimalizaci kvadratické funkce, tedy použít *minimalizační pravidlo*

$$\alpha_k \approx \arg \min_{\alpha > 0} f(x^k - \alpha g^k). \quad (4.4)$$

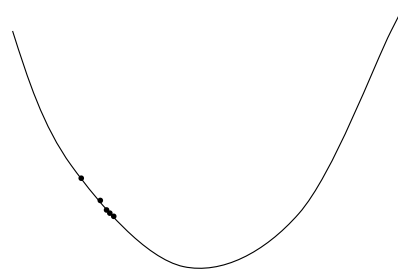
Snadnější je použít *minimalizační pravidlo s omezením*, tj. najít pro vhodně zvolené dané $s > 0$

$$\alpha_k \approx \arg \min_{\alpha \in (0, s)} f(x^k - \alpha g^k). \quad (4.5)$$

Při praktické implementaci, typicky realizované pomocí některé z metod popsaných v předchozí kapitole, nemůžeme očekávat, že se nám podaří najít minimum přesně, a musíme se spokojit se splněním nějakého volnějšího ukončovacího kritéria, které nám pokud možno zaručí alespoň konvergenci algoritmu. Pouhý pokles cenové funkce totiž nezaručuje konvergenci k lokálnímu minimu ani k stacionárnímu bodu. Intuitivně je třeba zabránit dlouhým krokům s malým poklesem, které jsou na obr. 4.3, a krátkým krokům s malým poklesem, které jsou na obr. 4.4. V žádném z uvedených příkladů nemusí být hromadné body vyobrazených posloupností stacionárními body f .



Obr. 4.3: Příliš dlouhý krok



Obr. 4.4: Příliš krátký krok

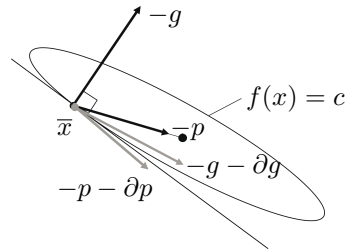
Nejčastěji se používá *Armijovo pravidlo*. Pravidlo používá tři pevně zvolené parametry $s > 0$, $0 < \beta < 1$, a $0 < \sigma < 1$. Délka kroku je rovna $\alpha_k = \beta^m s$, kde m je nejmenší nezáporné číslo, pro které

$$f(x^k) - f(x^k - \beta^m s d) \geq -\sigma \beta^m s \nabla f(x^k)^T d. \quad (4.6)$$

Tato podmínka zaručuje dostatečný pokles úměrný gradientu, čímž vylučuje oba nepříznivé případy. Podmínka se dá aplikovat i na jiné směry poklesu než je záporný gradient, například na Newtonův směr nebo jeho aproximace popsané v kapitole 6.

Robustnost gradientního směru

Důležitou vlastností gradientních metod je jejich robustnost. Z obr. 4.5 je patrné, že i když ∂g je relativně velká perturbace gradientu g , vektor $-g - \partial g$ je stále směrem poklesu, zatímco malá perturbace ∂p optimálního směru p může způsobit, že $-p - \partial p$ není ani směrem poklesu.



Obr. 4.5: Robustnost gradientu a směru maximálního poklesu g a p

4.3 Aproximace gradientu pomocí konečných diferencí

Některé důležité cenové funkce jsou diferencovatelné, avšak není možno zadat explicitní pravidlo pro výpočet gradientu. Jako příklad si uveďme typickou úlohu tvarové optimalizace v mechanice, kde $x \in \mathbb{R}^n$ jsou tzv. *návrhové proměnné* popisující tvar tělesa, ale cenový funkcionál f je zadán pomocí funkce

$$J : \mathbb{R}^{n_s} \rightarrow \mathbb{R}$$

definované pro tzv. *stavové proměnné* $u \in \mathbb{R}^{n_s}$, které splňují stavové rovnice

$$K(x)u = f(x).$$

V mechanice mohou být stavové proměnné posunutí a výše uvedená rovnice je rovnice rovnováhy definovaná pozitivně definitní maticí $K(x)$ a vektorem $f(x) \in \mathbb{R}^{n_s}$. K výpočtu hodnoty cenové funkce

$$f(x) = J(u(x))$$

pro dané x je tak třeba najít řešení $u = u(x)$ poměrně rozsáhlé (tisíce až miliony neznámých) soustavy rovnic rovnováhy. Návrhových proměnných bývá obvykle mnohem méně, řádově desítky až stovky.

Konečné diference

První derivace potřebné k vyčíslení gradientu mohou být nejsnáze vyčísleny pomocí *dopředných diferencí*

$$\frac{\partial f(x^k)}{\partial x_i} \approx \frac{1}{\varepsilon} (f(x^k + \varepsilon e_i) - f(x^k)), \quad (4.7)$$

kde e_i je i -tý sloupec jednotkové matice a $\varepsilon > 0$ je malé číslo. V přesné aritmetice by platilo čím menší, tím lepší, avšak při praktické počítačové implementaci je třeba toto pravidlo sladit s faktem, že čím menší ε , tím větší je vliv zaokrouhlovacích chyb způsobených počítačovou aritmetikou.

Dopředné diference jsou přesné pouze pro lineární funkce. Vyšší přesnosti je možno dosáhnout s *centrálními diferencemi*

$$\frac{\partial f(x^k)}{\partial x_i} \approx \frac{1}{2\varepsilon} (f(x^k + \varepsilon e_i) - f(x^k - \varepsilon e_i)), \quad (4.8)$$

kteřé jsou přesné pro kvadratické polynomy, avšak cena výpočtu je dvojnásobná. V řadě případů lze využít speciálního tvaru funkce k efektivnějšímu implicitnímu výpočtu hodnoty gradientu.

4.4 Ukončovací podmínky

Jako ukončovací podmínku používáme v této kapitole velikost gradientu, tj. $\|g^k\| = \|\nabla f(x^k)\|$. Tímto testem ověřujeme, zda je bod x^k dostatečně přesnou aproximací stacionárního bodu funkce f . Alternativní ukončovací podmínkou může být test na velikost kroku, tj. $\|x^{k+1} - x^k\|$. Jinou možností je ukončovat iterační minimalizační proces při dostatečně malé změně funkční hodnoty, tj. $\|f(x^{k+1}) - f(x^k)\|$.

Vhodnou ukončovací podmínkou je také test na relativní velikost gradientu.

Příklady k procvičení

1. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = x_1^2 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou největšího spádu s analyticky vypočteným gradientem a poté metodou největšího spádu s numericky vypočteným gradientem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

2. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = x_1^4 + x_1^2 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou největšího spádu s analyticky vypočteným gradientem a poté metodou největšího spádu s numericky vypočteným gradientem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

3. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100x_1^4 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou největšího spádu s analyticky vypočteným gradientem a poté metodou největšího spádu s numericky vypočteným gradientem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

4. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = \sin^2 x_1 + \cos^2 x_2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou největšího spádu s analyticky vypočteným gradientem a poté metodou největšího spádu s numericky vypočteným gradientem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

5. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = (x_1 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou největšího spádu s analyticky vypočteným gradientem a poté metodou největšího spádu s numericky vypočteným gradientem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

6. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = (x_1 - x_2^2)^2 + 100(1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou největšího spádu s analyticky vypočteným gradientem a poté metodou největšího spádu s numericky vypočteným gradientem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

7. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100(x_1^3 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou největšího spádu s analyticky vypočteným gradientem a poté metodou největšího spádu s numericky vypočteným gradientem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

8. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100(x_1 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou největšího spádu s analyticky vypočteným gradientem a poté metodou největšího spádu s numericky vypočteným gradientem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

Kapitola 5

Metoda sdružených gradientů

V této kapitole se budeme věnovat minimalizaci kvadratických funkcí, tj. zaměříme se na řešení úlohy

$$\min_{x \in \mathbb{R}^n} f(x), \quad (5.1)$$

kde $f(x) = \frac{1}{2}x^T A x - x^T b$, b je sloupcový vektor s n složkami a A je symetrická pozitivně definitní nebo pozitivně semidefinitní matice řádu $n \times n$. Zejména nás budou zajímat úlohy, kde n je velké a matice A je řídká a „dobře“ podmíněná. Je snadné ukázat, že řešení úlohy (5.1) je stejné jako řešení soustavy lineárních rovnic $Ax = b$.

K této minimalizaci budeme využívat gradientní informaci, ale důmyslnějším způsobem než v předchozí kapitole. Pro řešení uvedené úlohy zkonstruujeme tzv. metodu sdružených gradientů. Tato metoda při počítání v přesné aritmetice najde řešení po n krocích. Jde tedy o tzv. finitní metodu. Obvykle se ale metoda používá jako iterační a po k ($k \ll n$) krocích je výpočetní proces ukončen nalezením aproximace hledaného řešení.

Nakonec ukážeme, jak tuto metodu sdružených gradientů použít i pro minimalizaci funkcí, které nejsou kvadratické.

5.1 Metoda sdružených směrů pro kvadratické funkce

Metoda sdružených směrů je založena na několika jednoduchých pozorováních. Nyní provedeme první z nich, které nám umožňuje redukovat řešení úlohy (5.1) na řešení posloupnosti úloh jednorozměrné minimalizace.

Předpokládejme, že matice $A \in \mathbb{R}^{n \times n}$ je symetrická a pozitivně definitní a dále předpokládejme, že máme nenulové vektory p^1, \dots, p^n , pro které platí

$$(p^i, p^j)_A := (p^i)^T A p^j = 0 \quad \text{pro } i \neq j.$$

Takové vektory nazýváme *A-konjugované* nebo zkráceně *konjugované*. Pro tyto vektory platí, že jsou i lineárně nezávislé, jak snadno nyní ověříme. Jestliže máme ξ_1 až ξ_n , pro které platí

$$\xi_1 p^1 + \dots + \xi_n p^n = o,$$

pak přenásobením této rovnice zleva výrazem $(p^i)^T A$ a užitím konjugace dostaneme

$$\xi_1 (p^1)^T A p^1 + \dots + \xi_n (p^n)^T A p^n = \xi_i (p^i)^T A p^i = 0.$$

Protože matice A je symetrická pozitivně definitní, tak $(p^i)^T A p^i > 0$. Z toho vyplývá, že $\xi_i = 0$.

Z lineární nezávislosti vektorů p^1, \dots, p^n plyne, že tyto vektory tvoří bázi vektorového prostoru \mathbb{R}^n a libovolný vektor $x \in \mathbb{R}^n$ může být zapsán jako

$$x = \xi_1 p^1 + \dots + \xi_n p^n.$$

Substitucí do f a použitím konjugace dostaneme

$$\begin{aligned} f(x) &= \left(\frac{1}{2} \xi_1^2 (p^1)^T A p^1 - \xi_1 b^T p^1 \right) + \dots + \left(\frac{1}{2} \xi_n^2 (p^n)^T A p^n - \xi_n b^T p^n \right) \\ &= f(\xi_1 p^1) + \dots + f(\xi_n p^n). \end{aligned}$$

Z toho plyne

$$f(\bar{x}) = \min_{x \in \mathbb{R}^n} f(x) = \min_{\xi_1 \in \mathbb{R}} f(\xi_1 p^1) + \dots + \min_{\xi_n \in \mathbb{R}} f(\xi_n p^n).$$

Tím se nám povedlo převést původní úlohu (5.1) na n úloh jednorozměrné minimalizace. Protože

$$\left. \frac{df(\xi p^i)}{d\xi} \right|_{\xi_i} = \xi_i (p^i)^T A p^i - b^T p^i = 0,$$

pak řešení \bar{x} úlohy (5.1) je dáno předpisem

$$\bar{x} = \xi_1 p^1 + \dots + \xi_n p^n, \quad \xi_i = b^T p^i / (p^i)^T A p^i, \quad i = 1, \dots, n. \quad (5.2)$$

Pokud je n , které určuje dimenzi problému (5.1) „velké“, je úkol najít jeho řešení \bar{x} příliš náročný a drahý. V takovém případě je rozumné modifikovat schéma, které jsme popsali tak, abychom místo přesného řešení \bar{x} problému (5.1) obdrželi jeho aproximaci \tilde{x} s využitím počáteční aproximace x^0 a „několika“ vektorů p^1 až p^k , $k \ll n$. Přirozenou možností, jak získat aproximaci \tilde{x} je vektor x^k , ve kterém je dosaženo minima funkce f na množině $\mathcal{S}^k = x^0 + \langle p^1, \dots, p^k \rangle^1$. Abychom tuto aproximaci získali, všimněme si, že libovolný vektor $x \in \mathcal{S}^k$ lze zapsat ve tvaru

$$x = x^0 + \xi_1 p^1 + \dots + \xi_k p^k.$$

¹ $\langle v_1, \dots, v_k \rangle$ označuje lineární obal vektorů v_1, \dots, v_k

Po substituci do f a využitím toho, že p^1, \dots, p^k jsou konjugované, dostaneme

$$f(x) = f(x^0) + \left(\frac{1}{2} \xi_1^2 (p^1)^T A p^1 + \xi_1 (A x^0 - b)^T p^1 \right) + \dots \\ + \left(\frac{1}{2} \xi_k^2 (p^k)^T A p^k + \xi_k (A x^0 - b)^T p^k \right).$$

Označme reziduum $g^0 = g(x^0) = \nabla f(x^0) = Ax^0 - b$ a

$$f_0(x) = \frac{1}{2} x^T A x + x^T g^0.$$

Dostáváme

$$f(x) = f(x^0) + f_0(\xi_1 p^1) + \dots + f_0(\xi_k p^k)$$

a

$$f(x^k) = \min_{x \in \mathcal{S}^k} f(x) = f(x^0) + \min_{\xi_1 \in \mathbb{R}} f_0(\xi_1 p^1) + \dots + \min_{\xi_k \in \mathbb{R}} f_0(\xi_k p^k). \quad (5.3)$$

Tím jsme opět převedli naši úlohu na řešení posloupnosti jednoduchých jednorozměrných úloh. Aproximace řešení x^k je daná předpisem

$$x^k = x^0 + \xi_1 p^1 + \dots + \xi_k p^k, \quad \xi_i = -(g^0)^T p^i / (p^i)^T A p^i, \quad i = 1, \dots, k, \quad (5.4)$$

protože

$$\left. \frac{df(\xi p^i)}{d\xi} \right|_{\xi_i} = \xi_i (p^i)^T A p^i + (g^0)^T p^i = 0.$$

Protože podle (5.3) pro $k \geq 1$

$$f(x^k) = \min_{x \in \mathcal{S}^k} f(x) = f(x^{k-1}) + \min_{\xi \in \mathbb{R}} f_0(\xi p^k), \quad (5.5)$$

můžeme snadno generovat aproximace x^k iteračně. Metoda sdružených směrů začíná libovolnou počáteční aproximací x^0 . Jestliže aproximace x^{k-1} je daná, pak aproximací x^k získáme pomocí vztahu

$$x^k = x^{k-1} - \alpha_k p^k, \quad \alpha_k = (g^0)^T p^k / (p^k)^T A p^k. \quad (5.6)$$

Tzn. $f(x^{k-1} + \xi p^k)$ dosahuje svého minima v $\xi = -\alpha_k$ a rutina splňuje to, že postupné iterace x^k jsou minimy funkce f na množinách \mathcal{S}^k .

Pro koeficienty α_k můžeme odvodit alternativní vztah ke vztahu (5.6). Snadno lze ukázat, že z definice množiny \mathcal{S}^k plyne

$$(g^k)^T p^i = 0, \quad i = 1, \dots, k. \quad (5.7)$$

Protože pro $i \geq 1$ platí

$$g^i = Ax^i - b = A(x^{i-1} - \alpha_i p^i) - b = (Ax^{i-1} - b) - \alpha_i A p^i \\ = g^{i-1} - \alpha_i A p^i,$$

dostaneme pro $k \geq 1$ a $i = 1, \dots, k-1$, s využitím konjugace rovnost

$$(g^i)^T p^k = (g^{i-1})^T p^k - \alpha_i (p^i)^T A p^k = (g^{i-1})^T p^k.$$

Z toho plyne

$$(g^0)^T p^k = (g^1)^T p^k = \dots = (g^{k-1})^T p^k$$

a

$$\alpha_k = \frac{(g^0)^T p^k}{(p^k)^T A p^k} = \dots = \frac{(g^{k-1})^T p^k}{(p^k)^T A p^k}. \quad (5.8)$$

5.2 Metoda sdružených gradientů pro kvadratické funkce

Nyní se budeme zabývat problémem, jak efektivně získat konjugované vektory p^1 až p^n . Připomeňme si, jak vytvořit systém konjugovaných směrů pomocí *Gramm-Schmidtova procesu*. Předpokládejme, že p^1, \dots, p^k jsou nenulové konjugované směry, $1 \leq k < n$ a najděme způsob, jak použít vektor $h^k \notin \langle p^1, \dots, p^k \rangle$ tak, abychom obdrželi nový vektor p^{k+1} ve tvaru

$$p^{k+1} = h^k + \beta_{k1} p^1 + \dots + \beta_{kk} p^k. \quad (5.9)$$

Protože p^{k+1} má být konjugovaný k p^1, \dots, p^k , platí

$$\begin{aligned} 0 &= (p^i)^T A p^{k+1} = (p^i)^T A h^k + \beta_{k1} (p^i)^T A p^1 + \dots + \beta_{kk} (p^i)^T A p^k \\ &= (p^i)^T A h^k + \beta_{ki} (p^i)^T A p^i, \quad i = 1, \dots, k. \end{aligned}$$

Z toho plyne

$$\beta_{ki} = -\frac{(p^i)^T A h^k}{(p^i)^T A p^i}, \quad i = 1, \dots, k. \quad (5.10)$$

Zřejmě platí tvrzení

$$\langle p^1, \dots, p^{k+1} \rangle = \langle p^1, \dots, p^k, h^k \rangle.$$

Pokud máme lineárně nezávislé vektory h^0, \dots, h^{k-1} , můžeme zvolit $p^1 = h^0$ a použít vztahy (5.9) a (5.10) k získání množiny vzájemně A -konjugovaných směrů p^1 až p^k takových, že

$$\langle h^0, \dots, h^{i-1} \rangle = \langle p^1, \dots, p^i \rangle, \quad i = 1, \dots, k.$$

Pro obecnou množinu vektorů h^0, \dots, h^{k-1} je nalezení množiny A -konjugovaných směrů velmi náročné. Tato náročnost je způsobena jak paměťovými nároky, tak výpočetními nároky. Proto je vhodné zvolit jinou cestu. Touto alternativou je konstrukce konjugovaných směrů výše uvedenou procedurou, kterou aplikujeme na vektory z *Krylovových prostorů*

$$\mathcal{K}^k = \mathcal{K}^k(A, g^0) = \langle g^0, A g^0, \dots, A^{k-1} g^0 \rangle, \quad k = 1, \dots, n,$$

kde $g^0 = Ax^0 - b$, x^0 je vhodná počáteční aproximace a $\mathcal{K}^0 = \{o\}$. Tato efektivní metoda je opět založena na několika pozorováních.

Nejprve předpokládejme, že p^1, \dots, p^i tvoří konjugovanou bázi prostoru \mathcal{K}^i , $i = 1, \dots, k$, a všimněme si, že jestliže x^k označuje minimum funkce f na množině $x^0 + \mathcal{K}^k$, pak je gradient $g^k = \nabla f(x^k)$ ortogonální ke Krylovovu prostoru \mathcal{K}^k . Tzn.

$$(g^k)^T x = 0 \text{ pro libovolné } x \in \mathcal{K}^k.$$

Z toho přímo plyne, že jestliže $g^k \neq o$, potom

$$g^k \notin \mathcal{K}^k.$$

Protože platí $g^k \in \mathcal{K}^{k+1}$, můžeme použít (5.9), kde $h^k = g^k$, k rozšíření konjugované báze prostoru \mathcal{K}^k na konjugovanou bázi prostoru \mathcal{K}^{k+1} . Snadno ověříme, že platí

$$\mathcal{K}^k(A, g^0) = \langle g^0, \dots, g^{k-1} \rangle.$$

Dále si všimněme, že pro libovolné $x \in \mathcal{K}^{k-1}$ a $k \geq 1$ platí

$$Ax \in \mathcal{K}^k,$$

nebo stručně zapsáno $A\mathcal{K}^{k-1} \subseteq \mathcal{K}^k$. Protože $p^i \in \mathcal{K}^i \subseteq \mathcal{K}^{k-1}$, $i = 1, \dots, k-1$, pak platí

$$(Ap^i)^T g^k = (p^i)^T Ag^k = 0, \quad i = 1, \dots, k-1.$$

Z toho vyplývá, že

$$\beta_{ki} = -\frac{(p^i)^T Ag^k}{(p^i)^T Ap^i} = 0, \quad i = 1, \dots, k-1.$$

Výše uvedené využijeme pro následující odvození. Jestliže máme množinu takových konjugovaných vektorů p^1, \dots, p^k , pro které platí

$$\langle p^1, \dots, p^i \rangle = \mathcal{K}^i, \quad i = 1, \dots, k,$$

pak vztah (5.9) použitý na p^1, \dots, p^k a na $h^k = g^k$ se zredukuje na

$$p^{k+1} = g^k + \beta_k p^k, \quad (5.11)$$

kde

$$\beta_k = \beta_{kk} = -\frac{(p^k)^T Ag^k}{(p^k)^T Ap^k}. \quad (5.12)$$

Na závěr si všimněme, že z ortogonality g^k k prostoru $\langle p^1, \dots, p^k \rangle$ a vztahu (5.11) vyplývá, že

$$\|p^{k+1}\| \geq \|g^k\|. \quad (5.13)$$

A dále platí, že jestliže $g^{k-1} \neq o$, potom $p^k \neq o$. Tzn., že vztah (5.12) je dobře definován a navíc platí $g^{k-1} \neq o$.

V předcházejícím textu této kapitoly jsme zjistili, že konjugované směry můžeme s výhodou použít pro redukci minimalizace konvexní kvadratické n -rozměrné funkce na minimalizaci posloupnosti jednorozměrných funkcí. Dále jsme zjistili, že tyto konjugované směry mohou být generovány velmi efektivně pomocí Krylovových prostorů. Kombinací těchto dvou pozorování dostáváme velmi účinnou metodu - metodu sdružených gradientů.

Algoritmus začíná počáteční iterací x^0 , $g^0 = Ax^0 - b$, a $p^1 = g^0$. Jestliže máme dány x^{k-1} a g^{k-1} , kde $k \geq 1$, pak algoritmus nejprve ověří, zda x^{k-1} je řešení. Pokud nejde o řešení, pak algoritmus zkonstruuje další iteraci

$$x^k = x^{k-1} - \alpha_k p^k, \quad \text{kde } \alpha_k = (g^{k-1})^T p^k / (p^k)^T A p^k,$$

a

$$g^k = Ax^k - b = A(x^{k-1} - \alpha_k p^k) - b = (Ax^{k-1} - b) - \alpha_k A p^k = g^{k-1} - \alpha_k A p^k. \quad (5.14)$$

Nakonec je zkonstruován nový konjugovaný směr p^{k+1} pomocí vztahů (5.11) a (5.12).

Rozhodnutí, zda je x^{k-1} „dostatečně přesným“ řešením je obvykle učiněno na základě velikosti hodnoty $\|g^{k-1}\|$. Takže norma gradientu musí být vyčíslena v každé iteraci. Tato norma může být také použita k nahrazení skalárních součinů obsahujících gradient ve vztazích pro výpočet α_k a β_k . Abychom odvodili nové vztahy pro výpočet α_k a β_k , nahradíme k v (5.11) číslem $k-1$ a vynásobíme výsledný vztah vektorem $(g^{k-1})^T$. S využitím ortogonality získáme

$$(g^{k-1})^T p^k = \|g^{k-1}\|^2 + \beta_{k-1} (g^{k-1})^T p^{k-1} = \|g^{k-1}\|^2 \quad (5.15)$$

a po dosazení do (5.8) máme

$$\alpha_k = \frac{\|g^{k-1}\|^2}{(p^k)^T A p^k}. \quad (5.16)$$

Pro odvození vztahu pro β_k si všimněme, že $\alpha_k > 0$ pro $g^{k-1} \neq 0$ a ze vztahu (5.14) dostaneme

$$A p^k = -\frac{1}{\alpha_k} (g^k - g^{k-1}),$$

z čehož plyne

$$\alpha_k (g^k)^T A p^k = -(g^k)^T (g^k - g^{k-1}) = -\|g^k\|^2$$

a

$$\beta_k = -\frac{(p^k)^T A g^k}{(p^k)^T A p^k} = \frac{\|g^k\|^2}{\alpha_k (p^k)^T A p^k} = \frac{\|g^k\|^2}{\|g^{k-1}\|^2} = \frac{(g^k)^T (g^k - g^{k-1})}{\|g^{k-1}\|^2}. \quad (5.17)$$

Metoda sdružených gradientů je shrnuta v algoritmu 5.1.

Algoritmus 5.1 Metoda sdružených gradientů (CG).

Mějme symetrickou pozitivně definitní matici $A \in \mathbb{R}^{n \times n}$
a $b \in \mathbb{R}^n$.

Krok 0. {Inicializace.}

Zvolme $\varepsilon > 0$ (přesnost), $x^0 \in \mathbb{R}^n$, označme $g^0 = Ax^0 - b$,
 $p^1 = g^0$, $k = 1$

Krok 1. {Základní cyklus pro sdružené gradienty.}

while $\|g^{k-1}\| \geq \varepsilon$

$\alpha_k = \|g^{k-1}\|^2 / (p^k)^T A p^k$

$x^k = x^{k-1} - \alpha_k p^k$

$g^k = g^{k-1} - \alpha_k A p^k$

$\beta_k = \|g^k\|^2 / \|g^{k-1}\|^2 = \frac{(g^k)^T (g^k - g^{k-1})}{\|g^{k-1}\|^2}$

$p^{k+1} = g^k + \beta_k p^k$

$k = k + 1$

end while

Krok 2. {Řešení.}

$\bar{x} = x^k$

Poznámka 5.1. Každá iterace metody sdružených gradientů může být implementována tak, aby obsahovala pouze jedno násobení matice a vektoru. Toto násobení Hessiánem A je výpočetně nejnáročnější část jedné iterace. Do paměti se v průběhu výpočtu ukládají obvykle pouze vektory x^k , p^k , a g^k .

Připomeňme, že algoritmus najde v každé iteraci minimum x^k funkce f na množině $x^0 + \mathcal{K}^k = x^0 + \mathcal{K}^k(A, g^0)$ a rozšíří konjugovanou bázi prostoru \mathcal{K}^k na bázi prostoru \mathcal{K}^{k+1} tak, že $g^k \neq o$. Protože dimenze prostoru \mathcal{K}^k je menší nebo rovna k , vyplývá z toho, že pro nějaké $k \leq n$ platí

$$\mathcal{K}^k = \mathcal{K}^{k+1}.$$

Protože $g^k \in \mathcal{K}^{k+1}$ a g^k je ortogonální k \mathcal{K}^k , pak algoritmus 5.1, který je implementovaný v přesné aritmetice, najde řešení \bar{x} problému (5.1) v nejvýše n iteracích. Nejdůležitější vlastnosti algoritmu 5.1 můžeme shrnout v následující větě.

Věta 5.2. Necht $\{x^k\}$ je posloupnost generována algoritmem 5.1 pro nalezení řešení \bar{x} úlohy (5.1). Počáteční iterace je zvolena jako $x^0 \in \mathbb{R}^n$. Pak platí, že algoritmus je dobře definován a existuje $k \leq n$ takové, že $x^k = \bar{x}$. Navíc platí následující tvrzení pro $i = 1, \dots, k$:

- (i) $f(x^i) = \min\{f(x) : x \in x^0 + \mathcal{K}^i(A, g^0)\}$.
- (ii) $\|p^{i+1}\| \geq \|g^i\|$.
- (iii) $(g^i)^T g^j = 0$ pro $i \neq j$.
- (iv) $(p^i)^T A p^j = 0$ pro $i \neq j$.
- (v) $\mathcal{K}^i(A, g^0) = \langle g^0, \dots, g^{i-1} \rangle = \langle p^1, \dots, p^i \rangle$.

Poznámka 5.3. Obvykle je dostačující nalézt x^k takové, že číslo $\|g^k\|$ je malé. Například, pokud máme malé $\varepsilon > 0$, můžeme považovat $\|g^k\|$ za malé, pokud platí

$$\|g^k\| \leq \varepsilon \|b\|.$$

Potom je $\tilde{x} = x^k$ přibližné řešení splňující

$$\|A(\tilde{x} - \bar{x})\| \leq \varepsilon \|b\|, \quad \|\tilde{x} - \bar{x}\| \leq \varepsilon \lambda_{\min}(A)^{-1},$$

kde $\lambda_{\min}(A)$ označuje nejmenší vlastní číslo matice A . Snadno můžeme ověřit, že přibližné řešení \tilde{x} je řešením perturbované úlohy

$$\min_{x \in \mathbb{R}^n} \tilde{f}(x) = \min_{x \in \mathbb{R}^n} \frac{1}{2} x^T A x - \tilde{b}^T x, \quad \tilde{b} = b + g^k.$$

5.3 Rychlost konvergence

Uvedme nyní bez důkazu odhad rychlosti konvergence metody sdružených gradientů (pro kvadratické funkce).

Věta 5.4. Necht $\{x^k\}$ je posloupnost generována algoritmem 5.1 pro nalezení řešení \bar{x} úlohy (5.1). Počáteční iterace je zvolena jako $x^0 \in \mathbb{R}^n$. Potom chyba

$$e^k = x^k - \bar{x}$$

splňuje

$$\|e^k\|_A \leq 2 \left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right)^k \|e^0\|_A, \quad (5.18)$$

kde $\kappa(A) = \lambda_{\max}/\lambda_{\min}$ je číslo podmíněnosti matice A a $\|x\|_A^2 = x^T A x$ označuje energetickou normu x .

5.4 Metoda sdružených gradientů pro nelineární funkce

Metodu sdružených gradientů můžeme použít i pro obecnější funkce než kvadratické. Stačí pokud jsou tyto funkce diferencovatelné a jsme schopni najít nebo aspoň aproximovat jejich gradient. K aproximaci můžeme použít opět metodu konečných diferencí z předchozí kapitoly. Abychom mohli použít pro minimalizaci funkce, která není kvadratická, metodu sdružených gradientů, musíme učinit následující úpravy.

Vztah (5.16), který používáme pro určení délky kroku α_k v algoritmu 5.1 (toto α_k minimalizuje funkci f v bodě x^{k-1} ve směru $-p^k$), nahradíme rutinou provádějící jednorozměrnou minimalizaci funkce f v bodě x^{k-1} ve směru $-p^k$, tj. $\alpha_k = \arg \min_{\alpha > 0} f(x^{k-1} - \alpha p^k)$. K tomu můžeme využít například některý z algoritmů z kapitoly 3.

Dále musíme nahradit reziduum $g^k = g(x^k) = Ax^k - b = g^{k-1} - \alpha_k A p^k$, které je gradientem kvadratické funkce $f(x) = \frac{1}{2}x^T A x - x^T b$ v bodě x^k gradientem nelineární funkce f , tj. $g^k = \nabla f(x^k)$.

Pro výpočet β_k , které se v případě kvadratické funkce rovná $\frac{\|g^k\|^2}{\|g^{k-1}\|^2}$ i $\frac{(g^k)^T (g^k - g^{k-1})}{\|g^{k-1}\|^2}$, použijeme jeden z těchto vztahů. V případě nelineární funkce, ale tyto vztahy nedávají stejnou hodnotu β_k . V případě použití $\beta_k = \frac{\|g^k\|^2}{\|g^{k-1}\|^2}$ nazýváme tuto metodu metodou *Fletcher-Reeves*. Pokud použijeme vztah $\beta_k = \frac{(g^k)^T (g^k - g^{k-1})}{\|g^{k-1}\|^2}$ mluvíme o metodě *Polak-Ribiere*.

Algoritmus 5.2 Metoda nelineárních sdružených gradientů (CG).

```

Krok 0. {Inicializace.}
    Zvolme  $\varepsilon > 0$  (přesnost),  $x^0 \in \mathbb{R}^n$ , označme  $g^0 = \nabla f(x^0)$ ,
 $p^1 = g^0$ ,  $k = 1$ 
Krok 1. {Základní cyklus pro nelineární sdružené gradienty.}
    while  $\|g^{k-1}\| \geq \varepsilon$ 
         $\alpha_k = \arg \min_{\alpha > 0} f(x^{k-1} - \alpha p^k)$ 
         $x^k = x^{k-1} - \alpha_k p^k$ 
         $g^k = \nabla f(x^k)$ 
         $\beta_k = \|g^k\|^2 / \|g^{k-1}\|^2$  (Fletcher-Reeves)
         $\beta_k = (g^k)^T (g^k - g^{k-1}) / \|g^{k-1}\|^2$  (Polak-Ribiere)
         $p^{k+1} = g^k + \beta_k p^k$ 
         $k = k + 1$ 
    end while
Krok 2. {Řešení.}
     $\bar{x} = x^k$ 

```

Poznámka 5.5. Protože v případě použití metody nelineárních sdružených gradi-

entů už není zaručena konjugovanost vektorů p^1, \dots, p^n , může nastat problém, že vektor p^k je „téměř kolmý“ k vektoru $g^{k-1} = \nabla f(x^{k-1})$, tj. vektor $-p^k$ není směr poklesu a algoritmus 5.2 pak při linesearch $\alpha_k = \arg \min_{\alpha > 0} f(x^{k-1} - \alpha p^k)$ najde $\alpha_k \approx 0$. To znamená, že $g^k = \nabla f(x^k) \approx g^{k-1} = \nabla f(x^{k-1})$.

Pokud použijeme metodu Fletcher-Reeves (tj. $\beta_k = \|g^k\|^2 / \|g^{k-1}\|^2$), pak $\beta_k \approx 1$ a $p^{k+1} \approx g^k + p^k$. V případě, že $\|g^k\| \ll \|p^k\|$, pak $p^{k+1} \approx p^k$. V tom případě se algoritmus 5.2 „zacyklí“ v okolí bodu x^k a není schopen nalézt aproximaci minima funkce f .

Pokud použijeme metodu Polak-Ribiere (tj. $\beta_k = (g^k)^T (g^k - g^{k-1}) / \|g^{k-1}\|^2$), pak $\beta_k \approx 0$ a $p^{k+1} \approx g^k$. V tom případě je vektor $-p^{k+1} \approx -g^k$ směrem poklesu a algoritmus 5.2 provede v další iteraci jeden krok gradientní metody a začne znovu budovat konjugované vektory jako na začátku iteračního procesu metody nelineárních sdružených gradientů ($p^{k+1} \approx g^k$). V tomto případě algoritmus 5.2 pokračuje v hledání aproximace minima funkce f . Metoda Polak-Ribiere je proto stabilnější než metoda Fletcher-Reeves.

Příklady k procvičení

1. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = x_1^2 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou nelineárních sdružených gradientů.

2. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = x_1^4 + x_1^2 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou nelineárních sdružených gradientů.

3. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100x_1^4 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou nelineárních sdružených gradientů.

4. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = \sin^2 x_1 + \cos^2 x_2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou nelineárních sdružených gradientů.

5. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = (x_1 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou nelineárních sdružených gradientů.

6. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = (x_1 - x_2^2)^2 + 100(1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou nelineárních sdružených gradientů.

7. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100(x_1^3 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou nelineárních sdružených gradientů.

8. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100(x_1 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou nelineárních sdružených gradientů.

Kapitola 6

Newtonova metoda

V této kapitole se seznámíme s Newtonovou metodou, což je spádová metoda pro minimalizaci alespoň dvakrát spojitě diferencovatelné funkce f využívající kvadratický model f . Pokud je f alespoň dvakrát spojitě diferencovatelná a konvexní v okolí řešení, pak podle Taylorovy věty kvadratický model dobře aproximuje cenovou funkci v okolí řešení a můžeme očekávat velmi rychlou konvergenci. Cenou za to je vyšší cena iterace a menší robustnost Newtonova směru. Pokud není možno najít počáteční aproximaci blízko řešení, je třeba Newtonovu metodu modifikovat pomocí vhodné *globalizační strategie*.

6.1 Newtonova metoda

Newtonova metoda využívá kvadratický model f definovaný v okolí iterace x^k ,

$$f(x) \approx m(x) = m_k(x) = f(x^k) + (x - x^k)^T \nabla f(x^k) + \frac{1}{2} (x - x^k)^T \nabla^2 f(x^k) (x - x^k). \quad (6.1)$$

Minimum m je dosaženo v x , které řeší

$$\nabla m(x) = o, \quad \text{tzn.} \quad \nabla f(x^k) + \nabla^2 f(x^k)(x - x^k) = o.$$

Odtud dostáváme předpis pro Newtonovu iteraci

$$x^{k+1} = x^k - \nabla^2 f(x^k)^{-1} \nabla f(x^k). \quad (6.2)$$

Označíme-li si $d^k = -\nabla^2 f(x^k)^{-1} \nabla f(x^k)$, dostaneme pro pozitivně definitní Hessian $\nabla^2 f(x^k)$ a $\nabla f(x^k) \neq o$

$$\nabla f(x^k)^T d^k = -\nabla f(x^k)^T \nabla^2 f(x^k)^{-1} \nabla f(x^k) < 0,$$

takže Newtonův směr je směr poklesu a Newtonův krok lze považovat za krok varianty metody spádových směrů s pevně zvoleným krokem $\alpha_k = 1$. Poslední nerovnost plyne z pozorování, že inverzní matice k symetrické pozitivně definitní matici A je

pozitivně definitní, neboť každé $x \in \mathbb{R}^n$ lze napsat ve tvaru $x = Ay$ a pro symetrickou pozitivně definitní matici A platí

$$x^T A^{-1} x = (Ay)^T A^{-1} Ay = y^T Ay > 0.$$

Schéma Newtonovy metody vypadá následovně:

Algoritmus 6.1 Newtonova metoda.

Je dána dvakrát spojitě diferencovatelná funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

Krok 0. {Inicializace.}

Zvolme $\varepsilon > 0$ (přesnost), $x^0 \in \mathbb{R}^n$, polož $k = 0$

Krok 1. {Cyklus Newtonovy metody.}

while $\|\nabla f(x^k)\| \geq \varepsilon$

Najdi d^k , pro které platí $d^k = -\nabla^2 f(x^k)^{-1} \nabla f(x^k)$

$x^{k+1} = x^k + d^k$

$k = k + 1$

end while

Krok 2. {Dosad (přibližné) řešení.}

$\tilde{x} = x^k$

Protože výpočet $\nabla^2 f(x^k)^{-1}$ pro určení kroku d^k je velmi náročný, hledáme obvykle krok d^k jako řešení soustavy $\nabla^2 f(x^k) d^k = -\nabla f(x^k)$.

Použijeme-li k zajištění délky kroku Armijovo pravidlo (4.6) s počáteční délkou kroku $s = 1$, pak lze dokázat, že v blízkém okolí nesingulárního řešení není třeba dělat žádnou redukci. Tento výsledek není příliš překvapivý, neboť Newtonova metoda je založena na kvadratickém modelu funkce, který je blízko řešení věrný. Použijeme-li Newtonovu metodu k minimalizaci ryze konvexní kvadratické funkce, dostaneme řešení v jediném kroku.

Na druhé straně, pokud nejsme blízko řešení, tak Newtonův krok nemusí být ani směrem poklesu, neboť není zaručeno, že $\nabla^2 f(x^k)$ je pozitivně definitní. Newtonova metoda je proto nutné kombinovat s vhodnou globalizační strategií, která nahradí Newtonův krok nějakým krokem ve směru poklesu s přiměřenou délkou kroku.

6.2 Kvadratická rychlost konvergence

Jelikož Newtonova metoda používá věrnější model funkce než gradientní metoda, dá se očekávat, že její rychlost konvergence v okolí řešení bude větší než je rychlost konvergence gradientní metody nebo metody sdružených gradientů, alespoň pokud měříme rychlost konvergence počtem kroků a nezohledňujeme pracnost metod.

Věta 6.1. *Nechť f je dvakrát spojitě diferencovatelná funkce, která má minimum v \bar{x} a splňuje v okolí řešení*

$$d^T \nabla^2 f(x) d \geq \lambda \|d\|^2, \quad \lambda > 0.$$

Potom existuje okolí \bar{x} , ve kterém Newtonova metoda definovaná (6.2) konverguje superlineárně, tj. platí

$$\lim_{k \rightarrow \infty} \frac{\|x^{k+1} - \bar{x}\|}{\|x^k - \bar{x}\|} = \lim_{k \rightarrow \infty} \frac{o(\|x^{k+1} - \bar{x}\|)}{\|x^k - \bar{x}\|} = 0.$$

Důkaz. Označme si pro stručnost

$$g(x) = \nabla f(x), \quad x \in \mathbb{R}^n,$$

takže si můžeme přepsat Newtonův krok (6.2) do tvaru

$$x^{k+1} = x^k - \nabla g(x^k)^{-1} g(x^k). \quad (6.3)$$

Potom $g(\bar{x}) = 0$ a podle Taylorovy věty

$$g(\bar{x}) = g(x^k + (\bar{x} - x^k)) = g(x^k) + \nabla g(x^k)^T (\bar{x} - x^k) + o(\|x^k - \bar{x}\|).$$

Odtud dostaneme po přenásobení zleva $\nabla g(x^k)^{-1}$ a elementární úpravě

$$x^k - \bar{x} - \nabla g(x^k)^{-1} g(x^k) = o(\|x^k - \bar{x}\|).$$

Použijeme-li definici Newtonova kroku (6.3), dostaneme

$$x^{k+1} - \bar{x} = o(\|x^k - \bar{x}\|)$$

a

$$\frac{x^{k+1} - \bar{x}}{\|x^k - \bar{x}\|} = \frac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|}.$$

□

6.3 Newtonova metoda s modifikací Hessiánu

Z našich dosavadních úvah o Newtonově metodě vyplývá, že konverguje velmi rychle v okolí řešení, avšak obecně nemusí generovat ani směr poklesu, který bychom mohli použít k přiblížení k řešení. Naším prvním krokem k zajištění globální konvergence tedy bude zajistit, aby kvadratický model m_k funkce f v x^k byl ryze konvexní, tj.

aby Hessián H_k byl pozitivně definitní. Tím se současně zajistí, aby modifikovaná Newtonova metoda generovala směr poklesu, neboť pro pozitivně definitní matici H_k a $d = -H_k^{-1}\nabla f(x^k)$ platí

$$d^T \nabla f(x^k) = -\nabla f(x^k)^T H_k^{-1} \nabla f(x^k) < 0.$$

Pokud si nejsme jisti ryzí konvexitou cenové funkce, musíme tedy v každém kroku otestovat, zda je Hessián cenové funkce pozitivně definitní.

Násobek jednotkové matice

Je-li počet neznámých velmi malý, můžeme testovat pozitivní definitnost pomocí Sylvesterova pravidla. Pokud nahradíme v Newtonově kroku Hessián násobkem jednotkové matice cI , dostaneme směr poklesu gradientní metody a

$$x^{k+1} = x^k - c^{-1} I \nabla f(x^k),$$

tedy krok gradientní metody s fixní délkou kroku. Výsledkem je pomalá konvergence, což není překvapivé, neboť při tom nijak nevyužíváme informaci druhého řádu.

Další možností pro nalezení spektra matice H_k a tím ověření její pozitivní definitnosti je aplikovat na Hessián Geršgorinovu větu. Díky této větě získáme informaci o rozložení vlastních čísel Hessiánu a přičtením vhodného násobku jednotkové matice k Hessiánu získáme jeho pozitivně definitní aproximaci.

Choleského rozklad

Efektivní alternativou k prosté náhradě Hessiánu jinou pozitivně definitní maticí je využití Choleského rozkladu, který umožňuje rozložit pozitivně definitní A na součin

$$A = LL^T, \tag{6.4}$$

kde L je regulární dolní trojúhelníková matice s kladnými diagonálními prvky.

Metoda vyčíslení faktoru L se nazývá Choleského faktorizace. Choleského faktor L může být vypočten po sloupcích. Pokud

$$A = \begin{bmatrix} a_{11} & a_1^T \\ a_1 & A_{22} \end{bmatrix} \quad \text{a} \quad L = \begin{bmatrix} l_{11} & o \\ l_1 & L_{22} \end{bmatrix},$$

pak po dosazení za A a L do (6.4) a po porovnání odpovídajících prvků dostaneme

$$l_{11} = \sqrt{a_{11}}, \quad l_1 = l_{11}^{-1} a_1, \quad L_{22} L_{22}^T = A_{22} - l_1 l_1^T. \tag{6.5}$$

Tímto postupem získáme první sloupec matice L a zbývající blok L_{22} je Choleského faktor Schurova doplnku $A_{22} - l_1 l_1^T$. Protože je Schurův doplněk pozitivně definitní, můžeme najít jeho první sloupec výše uvedeným postupem atd. Tuto rutinu můžeme

implementovat tak, abychom využili řídký charakter matice A . Např. když $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ je *pásová matice*, kde $a_{ij} = 0$ pro $|i - j| > b$, $b \ll n$.

Pokud je matice $A \in \mathbb{R}^{n \times n}$ pouze pozitivně semidefinitní, může nastat, že $a_{11} = 0$. Potom

$$0 \leq x^T A x = y^T A_{22} y + 2x_1 a_1^T y$$

pro libovolný vektor $x = [x_1, y^T]^T$. Z nerovnosti vyplývá, že $a_1 = 0$, protože jinak bychom mohli zvolit $y = -a_1$ a velké x_1 takové, že

$$y^T A_{22} y + 2x_1 a_1^T y = a_1^T A_{22} a_1 - 2x_1 \|a_1\|^2 < 0.$$

Tedy pro symetrickou pozitivně definitní matici A a $a_{11} = 0$ se (6.5) redukuje na

$$l_{11} = 0, \quad l_1 = 0, \quad L_{22} L_{22}^T = A_{22}. \quad (6.6)$$

Aproximace Hessiánu

Hessián je možné aproximovat také numericky s využitím dopředných nebo centrálních diferencí. Stačí využít diferencí ze strany 28 a použít je pro aproximace druhých parciálních derivací.

Příklady k procvičení

1. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = x_1^2 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Newtonovou metodou s analyticky vypočteným gradientem i Hessiánem a poté Newtonovou metodou s numericky vypočteným gradientem i Hessiánem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

2. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = x_1^4 + x_1^2 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Newtonovou metodou s analyticky vypočteným gradientem i Hessiánem a poté Newtonovou metodou s numericky vypočteným gradientem i Hessiánem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

3. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100x_1^4 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Newtonovou metodou s analyticky vypočteným gradientem i Hessiánem a poté Newtonovou metodou s numericky vypočteným gradientem i Hessiánem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

4. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = \sin^2 x_1 + \cos^2 x_2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Newtonovou metodou s analyticky vypočteným gradientem i Hessiánem a poté Newtonovou metodou s numericky vypočteným gradientem i Hessiánem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

5. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = (x_1 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Newtonovou metodou s analyticky vypočteným gradientem i Hessiánem a poté Newtonovou metodou s numericky vypočteným gradientem i Hessiánem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

6. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = (x_1 - x_2^2)^2 + 100(1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Newtonovou metodou s analyticky vypočteným gradientem i Hessiánem a poté Newtonovou metodou s numericky vypočteným gradientem i Hessiánem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

7. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100(x_1^3 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Newtonovou metodou s analyticky vypočteným gradientem i Hessiánem a poté Newtonovou metodou s numericky vypočteným gradientem i Hessiánem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

8. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100(x_1 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Newtonovou metodou s analyticky vypočteným gradientem i Hessiánem a poté Newtonovou metodou s numericky vypočteným gradientem i Hessiánem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

Kapitola 7

Trust region

Jak jsme viděli výše, efektivnost Newtonovy metody pro minimalizaci hladké funkce f bez omezení záleží na věrnosti kvadratického modelu v okolí jeho minima. Dopusud jsme problém s případným rozdílem mezi cenovou funkcí a jejím lokálním modelem řešili zkrácením kroku. Alternativně však můžeme omezit platnost modelu také specifikací okolí Ω_k dané iterace x^k , na kterém je model věrný, a další iteraci volit podle

$$x^{k+1} = x^k + d^k, \quad d^k = \arg \min_{d \in \Omega_k} m_k(x^k + d).$$

Lze očekávat, že model funkce bude tímto způsobem využit lépe, než by tomu bylo u Newtonovy metody, která by zkrátila Newtonův krok tak, aby další iterace patřila do Ω_k , a to za cenu minimalizace komplikované omezení. Metody založené na minimalizaci modelu funkce na omezeném okolí dané iterace se anglicky nazývají metody typu *trust region*, tj. něco jako metody důvěryhodné oblasti.

Okolí Ω_k volíme tak, aby bylo možno minimum efektivně spočítat, nejčastěji pomocí ℓ_∞ -normy nebo pomocí eukleidovské normy, které jsou pro $y = [y_i] \in \mathbb{R}^n$ definovány po řadě předpisy

$$\|y\|_\infty = \max_{i=1,\dots,n} |y_i| \quad \text{a} \quad \|y\| = \sqrt{y_1^2 + \dots + y_n^2}.$$

V dalších kapitolách si ukážeme efektivní algoritmy pro hledání minima konvexní kvadratické funkce na okolí určeném ℓ_∞ -normou ve tvaru

$$\Omega_k = \{x \in \mathbb{R}^n : \|x - x^k\|_\infty \leq \delta_k\}.$$

V této kapitole se zaměříme na využití okolí určeného eukleidovskou normou

$$\Omega_k = \{x \in \mathbb{R}^n : \|x - x^k\|^2 \leq \delta_k\}.$$

Pro zjednodušení výkladu se omezíme na konvexní funkce, metodu však lze adaptovat na nalezení stacionárního bodu libovolné funkce.

7.1 Penalizační lemma

Náš výklad bude založen na výsledku, který nám umožní převést minimalizaci na eukleidovském okolí na minimalizaci modifikované cenové funkce bez omezení.

Lemma 7.1. *Nechť f je ryze konvexní funkce v \mathbb{R}^n , $x \in \mathbb{R}^n$, $\delta > 0$, a necht' $\bar{d} \in \mathbb{R}^n$ je řešením problému*

$$f(x + \bar{d}) = \min_{\|d\| \leq \delta} f(x + d). \quad (7.1)$$

Potom existuje $c > 0$ tak, že

$$\bar{d} = \arg \min_{d \in \mathbb{R}^n} \left(f(x + d) + \frac{c}{2} \|d\|^2 \right). \quad (7.2)$$

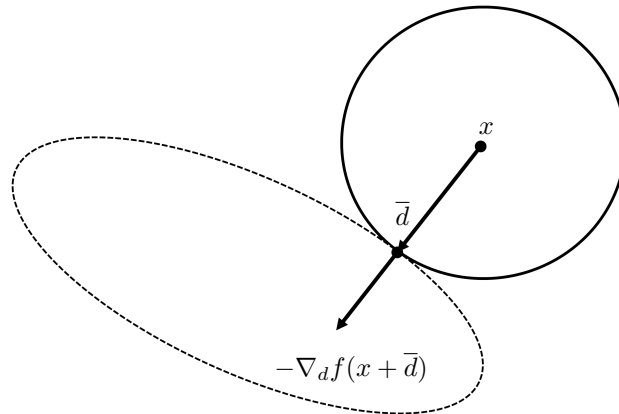
Důkaz. Jestliže řešení \bar{d} problému (7.1) splňuje $\|\bar{d}\| < \delta$, pak zřejmě stačí položit $c = 0$. Pokud tomu tak není, potom z obr. 7.1 plyne, že existuje $c > 0$ tak, že platí

$$c\bar{d} = -\nabla_d f(x + \bar{d}).$$

Odtud

$$\nabla_d \left(f(x + d) + \frac{c}{2} \|d\|^2 \right) \Big|_{d=\bar{d}} = \nabla_d f(x + \bar{d}) + c\bar{d} = 0,$$

což je podmínka pro minimum funkce na pravé straně rovnosti (7.2).



Obr. 7.1: Ilustrace penalizačního lemmatu

□

7.2 Levenberg-Marquardtova implementace

Budeme používat kvadratický model \overline{m}_k funkce f v okolí iterace x^k , který zahrnuje omezení minimalizace na okolí Ω_k pomocí $\delta_k \geq 0$, takže

$$\overline{m}_k(x^k + d) = f(x^k) + d^T \nabla f(x^k) + \frac{1}{2} d^T (\nabla^2 f(x^k) + c_k I) d. \quad (7.3)$$

Algoritmus nejprve najde „zkušební“ délku kroku d_k přibližným řešením

$$d^k \approx \arg \min_{d \in \mathbb{R}^n} \overline{m}_k(x^k + d).$$

Tento krok odpovídá podle lemmatu 7.1 minimalizaci m_k na nějakém okolí Ω_k určeném implicitně δ_k . Pak se rozhodne, zda ho přijme podle toho, zda se podařilo dosáhnout dostatečného poklesu cenové funkce. V dalším kroku se aktualizuje δ_k , tj. velikost okolí, na kterém se provádí minimalizace. Celý algoritmus potom vypadá následovně:

Algoritmus 7.1 Základní Trust Region algoritmus.

Je dána diferencovatelná funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

Krok 0. {Inicializace.}

Zvolme $\varepsilon > 0$ (přesnost), $0 < \eta_1 \leq \eta_2 < 1$, $0 < \gamma_1 < 1 < \gamma_2$, x^0 , $k = 0$, $\delta_0 > 0$.

Krok 1. {Cyklus gradientní metody.}

while $\|\nabla f(x^k)\| \geq \varepsilon$

Krok 2. {Minimalizace modelu \overline{m}_k .}

$d^k \approx \arg \min_{d \in \mathbb{R}^n} \overline{m}_k(x^k + d)$

Krok 3. {Aktualizace x^k .}

$\rho_k = \frac{f(x^k) - f(x^k + d^k)}{m_k(x^k) - \overline{m}_k(x^k + d^k)}$

if $\rho_k \geq \eta_1$

then $x^{k+1} = x^k + d^k$

else $x^{k+1} = x^k$

end if

Krok 4. {Trust region aktualizace (pomocí aktualizace c_k).}

$$c_k = \begin{cases} \gamma_1 c_k & \text{if } \rho_k > \eta_2 \\ c_k & \text{if } \rho_k \in \langle \eta_1, \eta_2 \rangle \\ \gamma_2 c_k & \text{if } \rho_k < \eta_1 \end{cases}$$

$k = k + 1$

end while

Krok 5. {Dosad (přibližné) řešení.}

$\tilde{x} = x^k$

Algoritmus 7.1 lze považovat též za prototyp algoritmu tzv. sekvenčního kvadratického programování (sequential quadratic programming - SQP). Modifikací pravidla aktualizace δ_k ho lze snadno upravit tak, aby hledal stacionární bod i pro funkce, které nejsou konvexní. Zde jsme se omezili na konvexní funkce částečně pro jednoduchost, částečně proto, abychom se vyhnuli použití lemmatu 7.1 pro nekonvexní funkce. Poznamenejme, že lemma 7.1 platí i pro nekonvexní funkce, její důkaz je téměř triviální aplikace podmínky minima funkce s omezením, kterou si však odvodíme až v dalších kapitolách.

Příklady k procvičení

1. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = x_1^2 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Trust region metodou. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

2. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = x_1^4 + x_1^2 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Trust region metodou. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

3. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100x_1^4 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Trust region metodou. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

4. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = \sin^2 x_1 + \cos^2 x_2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Trust region metodou. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

5. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = (x_1 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Trust region metodou. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

6. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = (x_1 - x_2^2)^2 + 100(1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Trust region metodou. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

7. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100(x_1^3 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte Trust region metodou. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

8. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100(x_1 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

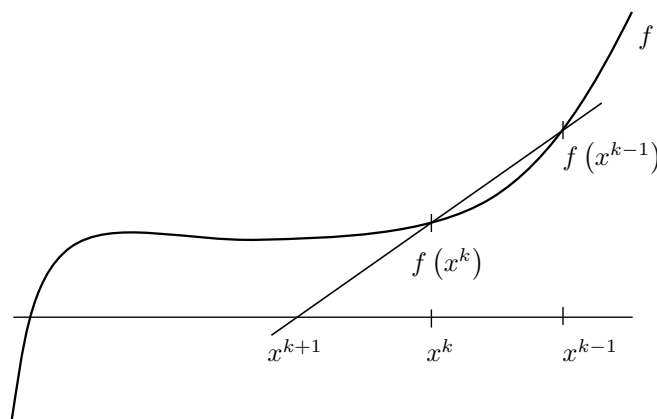
Tuto úlohu řešte Trust region metodou. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

Kapitola 8

Metody aproximace Hessiánu - BFGS

8.1 Heuristické odvození základních myšlenek metody BFGS

V této kapitole vybudujeme minimalizační algoritmus pro metodu, která používá Newtonův krok. Ale narozdíl od Newtonovy metody je informace druhého řádu (tj. v případě funkce jedné proměnné druhá derivace a v případě funkce více proměnných Hessián) aproximována na základě znalostí předcházejících iterací. Takové metodě budeme říkat *kvazi-Newtonova*. K této aproximaci informace druhého řádu použijeme v případě funkce jedné proměnné tzv. *metodu sečen* (viz obrázek 8.1) určenou pro hledání řešení rovnice $f(x) = 0$. Tato metoda aproximuje řešení průsečíkem sečny



Obr. 8.1: Algoritmus metody sečen

ke grafu funkce f v bodech x^k a x^{k-1} s osou x , tzn. $x^{k+1} = x^k - f(x^k) \frac{x^k - x^{k-1}}{f(x^k) - f(x^{k-1})}$

a $\frac{f(x^k) - f(x^{k-1})}{x^k - x^{k-1}} \approx f'(x^k)$. Pokud použijeme metodu sečen pro řešení rovnice $f'(x) = 0$, dostaneme vztah

$$f'(x^k) - f'(x^{k-1}) = b(x^k - x^{k-1}), \quad (8.1)$$

který použijeme pro aproximaci druhé derivace v bodě x^k ($b \approx f''(x^k)$). V případě funkce více proměnných použijeme zobecnění vztahu (8.1):

$$\nabla f(x^k) - \nabla f(x^{k-1}) = B(x^k - x^{k-1}), \quad (8.2)$$

který použijeme pro aproximaci Hessiánu v bodě x^k ($B \approx \nabla^2 f(x^k)$). Tento vztah již nestačí pro jednoznačné určení matice B a proto musíme k jejímu určení přidat další podmínku. Tu nám poskytne lemma uvedená v další podkapitole.

8.2 Minimalizační lemma

Lemma 8.1. *Nechť $y, s \in \mathbb{R}^n$ a množina $Q(y, s)$ je dána $Q(y, s) = \{B \in \mathbb{R}^{n \times n} : Bs = y\}$. Pak platí:*

$$\min_{B \in Q(y, s)} \|B - A\| = \|A_+ - A\|, \quad (8.3)$$

kde

$$A_+ = A + \frac{(y - As)s^T}{s^T s}, \quad (8.4)$$

$A_+ \in Q(y, s)$ a $A \in \mathbb{R}^{n \times n}$.

Důkaz. Ukážeme nejprve, že platí vztah (8.3), tzn. $\forall B \in Q(y, s)$:

$$\|A_+ - A\| \leq \|B - A\|.$$

K tomu použijeme vztah (8.4), definici množiny $Q(y, s)$ a vlastnost normy $\|MN\| \leq \|M\|\|N\|$, která platí pro všechny matice $M, N \in \mathbb{R}^{n \times n}$.

$$\begin{aligned} \text{Platí} \\ \|A_+ - A\| &= \left\| A + \frac{(y - As)s^T}{s^T s} - A \right\| = \left\| \frac{(y - As)s^T}{s^T s} \right\| = \left\| \frac{(Bs - As)s^T}{s^T s} \right\| = \left\| (B - A) \frac{ss^T}{s^T s} \right\| \leq \\ &\leq \|B - A\| \frac{\|ss^T\|}{s^T s}. \end{aligned}$$

Pokud využijeme toho, že norma matice je indukovaná vektorovou normou a tedy platí

$$\|M\| = \max \left\{ \frac{\|Mx\|}{\|x\|} : x \in \mathbb{R}^n, x \neq o \right\},$$

pak je snadné dokázat, že platí

$$\frac{\|ss^T\|}{s^T s} = 1.$$

Z toho plyne

$$\|B - A\| \frac{\|ss^T\|}{s^T s} = \|B - A\|$$

a tím jsme ověřili platnost $\forall B \in Q(y, s)$:

$$\|A_+ - A\| \leq \|B - A\|.$$

Je snadné ověřit, že platí $A_+ \in Q(y, s)$. Tím jsme ukázali, že $\|B - A\|$ nabývá svého minima na $Q(y, s)$ v A_+ , které je dáno vztahem

$$A_+ = A + \frac{(y - As)s^T}{s^T s}.$$

□

8.3 Symetrická pozitivně definitní aproximace Hessiánu

V této části vybudujeme symetrickou pozitivně definitní aproximaci Hessiánu na základě znalostí gradientů funkce f z předcházejících iterací. V případě funkce jedné proměnné platí $f'(x+h) - f'(x) = f''(\xi)h \approx f''(x)h$. Pokud tuto rovnost zobecníme do více dimenzí, dostaneme

$$H^k s = y, \tag{8.5}$$

kde

$$s = x^k - x^{k-1}, \quad y = \nabla f(x^k) - \nabla f(x^{k-1}).$$

Předpokládejme, že máme k dispozici Choleského rozklad matic H^k a H^{k-1} :

$$H^k = JJ^T, \tag{8.6}$$

$$H^{k-1} = LL^T.$$

Dosaďme vztah (8.6) do vztahu (8.5), tím dostaneme

$$JJ^T s = y.$$

Označme $v = J^T s$ a tím dostaneme $Jv = y$.

Zvolme matici J tak, aby splňovala $Jv = y$ a současně byla její vzdálenost od matice L minimální. K tomu můžeme použít tvrzení lematu 8.1. Dostaneme

$$J = L + \frac{(y - Lv)v^T}{v^T v}. \tag{8.7}$$

V rovnici $v = J^T s$ dosaďme za J ze vztahu (8.7):

$$L^T s + \frac{v(y - Lv)^T s}{v^T v} = v. \tag{8.8}$$

Dále upravme (8.8) a dostaneme

$$L^T s + \frac{vy^T s - vv^T L^T s}{v^T v} = v$$

a po další úpravě

$$L^T s + \frac{y^T s - v^T L^T s}{v^T v} v = v. \quad (8.9)$$

Označme v (8.9) zlomek konstantou k :

$$k = \frac{y^T s - v^T L^T s}{v^T v}.$$

A upravme (8.9) a získáme

$$L^T s = (1 - k)v. \quad (8.10)$$

Pokud označíme $\frac{1}{1-k}$ jako α , prepíšeme vztah (8.10) na

$$\alpha L^T s = v. \quad (8.11)$$

Dosaďme do (8.8) za v ze vztahu (8.11) a dostaneme

$$L^T s + \alpha \frac{y^T s - \alpha s^T L L^T s}{\alpha^2 s^T L L^T s} L^T s = \alpha L^T s$$

a po úpravě

$$L^T s + \alpha \frac{y^T s L^T s}{\alpha^2 s^T L L^T s} - L^T s = \alpha L^T s.$$

Tím získáme

$$\alpha^2 = \frac{y^T s}{s^T L L^T s} = \frac{y^T s}{s^T H^{k-1} s}.$$

V případě, že je splněna podmínka $y^T s > 0$ ¹, pak vyjádříme α :

$$\alpha = \sqrt{\frac{y^T s}{s^T H^{k-1} s}}.$$

Pokud podmínka $y^T s > 0$ není splněna, pak je třeba zvolit $H^k = H^{k-1}$, abychom měli nadále zaručenu pozitivní definitnost aproximace Hessiánu.

Po dosazení do (8.11):

$$v = \sqrt{\frac{y^T s}{s^T H^{k-1} s}} L^T s. \quad (8.12)$$

¹Tato podmínka je splněna, pokud je funkce f konvexní v bodě x^k , neboť z toho plyne, že matice H^k je symetrická pozitivně definitní, a proto $y^T s = s^T H^k s = s^T H^{k-1} s > 0$. V případě, že funkce f není konvexní, pak je podmínka splněna, pokud pro délku kroku αs navíc vyžadujeme tzv. *Wolfeho podmínku* $\nabla f(x^{k-1} + \alpha s)^T s \geq \beta \nabla f(x^{k-1})^T s$, $0 < \beta < 1$. A to proto, že $y^T s = \nabla f(x^k)^T s - \nabla f(x^{k-1})^T s$ a $\nabla f(x^{k-1})^T s < 0$ platí, protože s je směr poklesu.

Dosadme do (8.6) vztah (8.7):

$$H^k = JJ^T = \left(L + \frac{(y - Lv)v^T}{v^T v} \right) \left(L + \frac{(y - Lv)v^T}{v^T v} \right)^T = LL^T + \frac{yy^T}{v^T v} - \frac{Lvv^T L^T}{v^T v}. \quad (8.13)$$

Použijeme (8.12) a odvodíme

$$v^T v = s^T LL^T s \frac{y^T s}{s^T H^{k-1} s} = s^T H^{k-1} s \frac{y^T s}{s^T H^{k-1} s} = y^T s \Rightarrow \frac{yy^T}{v^T v} = \frac{yy^T}{y^T s} \quad (8.14)$$

a

$$vv^T = \frac{y^T s}{s^T H^{k-1} s} L^T s s^T L \Rightarrow \frac{Lvv^T L^T}{v^T v} = \frac{y^T s}{s^T H^{k-1} s} \frac{LL^T s s^T LL^T}{y^T s} = \frac{H^{k-1} s s^T H^{k-1}}{s^T H^{k-1} s}. \quad (8.15)$$

Po dosazení (8.14) a (8.15) do (8.13) získáme vztah pro symetrickou pozitivně definitní aproximaci Hessiánu v bodě x^k :

$$H^k = H^{k-1} + \frac{yy^T}{y^T s} - \frac{H^{k-1} s s^T H^{k-1}}{s^T H^{k-1} s}. \quad (8.16)$$

8.4 Algoritmus BFGS

Vztah (8.16) je základním kamenem metody BFGS (Broyden-Fletcher-Goldfarb-Shanno). Při budování symetrické pozitivně definitní aproximace Hessiánu pomocí tohoto vztahu je rozumné zvolit jako H^0 jednotkovou matici I .

Algoritmus 8.1 Metoda BFGS.

Krok 0. {Inicializace.}
 Zvolme $\varepsilon > 0$ (přesnost), $x^0 \in \mathbb{R}^n$, označme $g^0 = \nabla f(x^0)$, $H^0 = I$ (jednotková matice), $k = 1$

Krok 1. {Základní cyklus pro metodu BFGS.}
while $\|g^{k-1}\| \geq \varepsilon$
 $x^k = x^{k-1} - (H^{k-1})^{-1} g^{k-1}$ (Newtonův krok)
 $g^k = \nabla f(x^k)$
 $y = g^k - g^{k-1}$
 $s = x^k - x^{k-1}$
if $y^T s > 0$

$$H^k = H^{k-1} + \frac{yy^T}{y^T s} - \frac{H^{k-1} s s^T H^{k-1}}{s^T H^{k-1} s}$$
 (symetrická pozitivně definitní aproximace Hessiánu)
else
 $H^k = H^{k-1}$
end
 $k = k + 1$
end while

Krok 2. {Řešení.}
 $\bar{x} = x^k$

Protože výpočet $(H^{k-1})^{-1}$ je velmi náročný, nahrazujeme obvykle výpočet $-(H^{k-1})^{-1} g^{k-1}$ řešením soustavy $(H^{k-1}) d = -g^{k-1}$.

Příklady k procvičení

1. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = x_1^2 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou BFGS. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

2. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = x_1^4 + x_1^2 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou BFGS. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

3. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100x_1^4 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou BFGS. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

4. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = \sin^2 x_1 + \cos^2 x_2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou BFGS. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

5. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = (x_1 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou BFGS. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

6. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = (x_1 - x_2^2)^2 + 100(1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou BFGS. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

7. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100(x_1^3 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou BFGS. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

8. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = 100(x_1 - x_2^2)^2 + (1 - x_1)^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou BFGS. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.

Kapitola 9

Podmínky minima pro úlohy s omezením ve tvaru rovností

Počínaje touto kapitolou se budeme zabývat řešením optimalizačních úloh s omezením. V technických úlohách takové omezení vzniká například potřebou zahrnout do formulace úlohy omezení zdrojů, které máme k dispozici, nebo požadavky na pevnost konstrukcí. Začneme s omezeními ve tvaru rovností a ukážeme si, jak poznat řešení úlohy

$$\min_{x \in \Omega} f(x), \quad \Omega = \{x \in \mathbb{R}^n : h(x) = o\}, \quad h : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad m < n. \quad (9.1)$$

Budeme při tom předpokládat, že f je dvakrát spojitě diferencovatelná funkce a h je spojitě diferencovatelná. Množina Ω se nazývá přípustná množina.

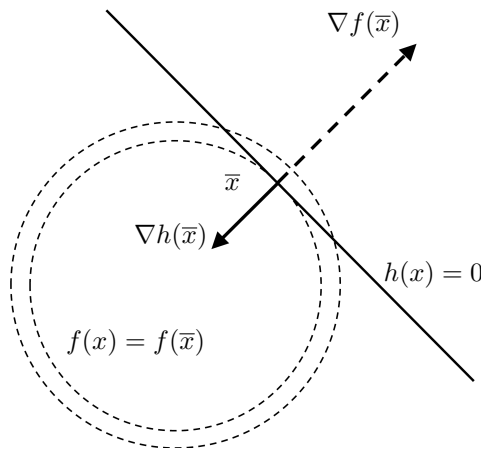
Podmínky minima pro funkce s omezením mají význam pro zodpovězení základních teoretických otázek optimalizace. Jejich znalost je užitečná i mimo optimalizaci. Uplatní se například při studiu fyziky a technických aplikací. I když je „ruční“ řešení akademických úloh velmi často založeno na řešení rovnic (a nerovnic), které formálně popisují podmínky minima, počítačové řešení rozsáhlejších praktických úloh je obvykle založeno na jiných postupech, popsanych v dalších kapitolách. Úlohy s omezením lze řešit bez počítače jen výjimečně.

9.1 Geometrická ilustrace podmínky minima

Začneme s řešením jednoduché úlohy

$$\min_{h(x)=0} f(x), \quad f(x) = x_1^2 + x_2^2, \quad h(x) = 4 - x_1 - x_2. \quad (9.2)$$

Pomůžeme si obrázkem 9.1, na kterém jsou znázorněny čárkovaně kruhové vrstevnice funkce f a nepřerušovanou čarou je znázorněna přímka, která představuje přípustnou množinu. Minimum je dosaženo v bodě \bar{x} .



Obr. 9.1: Minimum kvadratické funkce na přímce

Jelikož minimum f bez omezení je dosaženo ve společném středu všech kružnic popsanych pomocí rovnic $f(x) = c$, $c > 0$, je z obrázku vidět, že nejmenší z kružnic, na nichž leží alespoň jeden přípustný bod, je právě ta kružnice, která se dotýká přípustné množiny. Ve společném bodě \bar{x} platí, že gradient f je násobkem gradientu h , tedy existuje $\lambda \in \mathbb{R}$ tak, že

$$\nabla f(\bar{x}) + \lambda \nabla h(\bar{x}) = o.$$

Dostali jsme tak algebraickou formu našeho pozorování. Jelikož řešení musí být přípustné, můžeme napsat rovnice, které v našem případě určují řešení:

$$\nabla f(x) + \lambda \nabla h(x) = o, \quad h(x) = o.$$

Jelikož první rovnice je vektorová, můžeme si je přepsat jako tři skalární rovnice

$$2x_1 - \lambda = 0, \quad 2x_2 - \lambda = 0, \quad x_1 + x_2 = 4.$$

Jejich řešením dostaneme složky řešení $\bar{x}_1 = \bar{x} = 2$, $\lambda = 4$. Povšimněte si, že jsme získali nejen řešení úlohy (9.2), ale také hodnotu λ , na kterou se nás nikdo neptal. V dalším výkladu uvidíme, že λ je důležitá veličina, jejíž pochopení nám umožní zavést nové algoritmy a dodatečnou informaci o kvalitě řešení.

9.2 Úloha kvadratického programování s rovnostmi a Lagrangián

Odvodíme si nyní podmínky minima pro úlohu konvexního kvadratického programování

$$\min_{x \in \Omega} f(x), \quad f(x) = \frac{1}{2} x^T A x - b^T x, \quad \Omega = \{x \in \mathbb{R}^n : Bx = c\}, \quad (9.3)$$

kde A je symetrická pozitivně definitní matice řádu n , B je typu $m \times n$ a $c \in \mathcal{H}(B)$ ¹. Předpokládáme, že $B \neq O$ a že hodnost B je různá od n , takže nulový prostor $\mathcal{N}(B) \neq \{o\}$ ², avšak nevyžadujeme, aby B měla nezávislé řádky. Je snadné ukázat, že Ω je neprázdná uzavřená konvexní množina.

Přípustná množina Ω je *afinní prostor* tvaru

$$\Omega = \bar{x} + \mathcal{N}(B),$$

kde \bar{x} je jakýkoliv vektor, který splňuje

$$B\bar{x} = c.$$

Provedeme-li substituci $x = \bar{x} + z$, $z \in \mathcal{N}(B)$, můžeme převést (9.3) na minimalizaci

$$f_{\bar{x}}(z) = \frac{1}{2}z^T A z - (b - A\bar{x})^T z \quad (9.4)$$

na lineárním prostoru $\mathcal{N}(B)$. Když se nám to bude hodit, můžeme proto předpokládat, že $c = o$ v definici Ω .

Užitečným nástrojem k analýze a řešení problémů s omezením je *Lagrangián* $L_0 : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ definovaný pomocí

$$L_0(x, \lambda) = f(x) + \lambda^T (Bx - c) = \frac{1}{2}x^T A x - b^T x + (Bx - c)^T \lambda. \quad (9.5)$$

Povšimněme si, že se nám gradient Lagrangiánu objevil při algebraické formulaci podmínek minima v našem úvodním příkladu. Zřejmě

$$\nabla_{xx}^2 L_0(x, \lambda) = \nabla^2 f(x) = A, \quad (9.6)$$

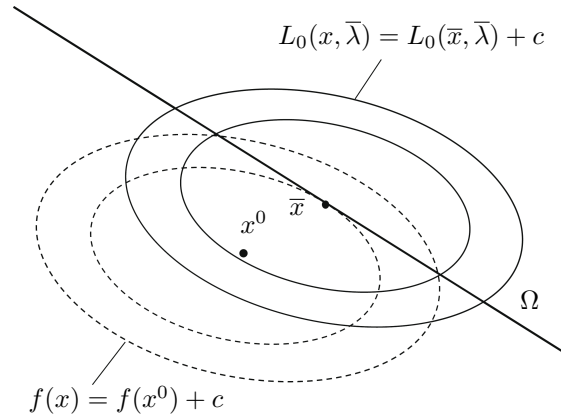
$$\nabla_x L_0(x, \lambda) = \nabla f(x) + B^T \lambda = Ax - b + B^T \lambda, \quad (9.7)$$

$$L_0(x + d, \lambda) = L_0(x, \lambda) + (Ax - b + B^T \lambda)^T d + \frac{1}{2}d^T A d. \quad (9.8)$$

Lagrangián je v případě kvadratického programování definován tak, že pokud se na něho díváme jako na funkci x , potom je jeho Hessián a restrikce na Ω identická s f , avšak jeho gradient $\nabla_x L_0(x, \lambda)$ závisí na hodnotě λ . Odtud plyne, že je-li f konvexní, potom L_0 je konvexní pro jakékoliv pevně zvolené λ , a globální minimum L_0 vzhledem k x se rovněž mění s λ . Uvidíme, že je možné zformulovat podmínky pro A , B , a b tak, že s vhodnou volbou $\lambda = \bar{\lambda}$ se řešení problému (9.3) s omezením redukuje na minimalizaci L_0 bez omezení tak jako na obr. 9.2.

¹ $\mathcal{H}(B)$ označuje obor hodnot matice B .

² $\mathcal{N}(B)$ označuje nulový prostor (jádro) matice B .



Obr. 9.2: Geometrický význam Lagrangianu

9.3 Podmínky minima pro kvadratické programování s rovnostmi

Hlavní otázky týkající se podmínek optimality a řešitelnosti (9.3) jsou zformulovány v následujícím tvrzení.

Věta 9.1. *Nechť je problém (9.3) definován pomocí symetrické matice $A \in \mathbb{R}^{n \times n}$, matice omezení $B \in \mathbb{R}^{m \times n}$ se sloupcovou hodnotí menší než n , a vektorů $b \in \mathbb{R}^m$, $c \in \mathcal{H}(B)$. Potom platí:*

(i) *Vektor $\bar{x} \in \Omega$ je řešení (9.3), právě když $A|_{\mathcal{N}(B)}$ je pozitivně semidefinitní a*

$$(A\bar{x} - b)^T d = 0 \quad (9.9)$$

pro každé $d \in \mathcal{N}(B)$.

(ii) *Vektor $\bar{x} \in \Omega$ je řešení (9.3), právě když $A|_{\mathcal{N}(B)}$ je pozitivně semidefinitní a existuje vektor $\bar{\lambda} \in \mathbb{R}^m$ tak, že*

$$A\bar{x} - b + B^T \bar{\lambda} = o. \quad (9.10)$$

Důkaz. (i) Nechť \bar{x} je řešení problému (9.3), takže pro každé $d \in \mathcal{N}(B)$ a $\alpha \in \mathbb{R}$

$$0 \leq f(\bar{x} + \alpha d) - f(\bar{x}) = \alpha(A\bar{x} - b)^T d + \frac{\alpha^2}{2} d^T A d. \quad (9.11)$$

Zvolíme-li $d \in \mathcal{N}(B)$ a α dostatečně velké, dostaneme z nezápornosti pravé strany (9.11), že $d^T A d \geq 0$. Tedy $A|_{\mathcal{N}(B)}$ musí být pozitivně semidefinitní. Na druhé straně, pro dostatečně malé hodnoty α a $(A\bar{x} - b)^T d \neq 0$, znaménko pravé strany rovnosti (9.11) je určeno znaménkem $\alpha(A\bar{x} - b)^T d$. Jelikož si můžeme zvolit znaménko

α libovolně a pravá strana (9.11) je nezáporná, dojdeme k závěru že (9.9) platí pro každé $d \in \mathcal{N}(B)$.

Předpokládejme nyní, že (9.9) platí pro vektor $\bar{x} \in \Omega$ a že $A|\mathcal{N}(B)$ je pozitivně semidefinitní. Potom

$$f(\bar{x} + d) - f(\bar{x}) = \frac{1}{2}d^T Ad \geq 0$$

pro každé $d \in \mathcal{N}(B)$, takže \bar{x} je řešení (9.3).

(ii) Necht \bar{x} je řešení (9.3), takže podle (i) $A|\mathcal{N}(B)$ je pozitivně semidefinitní a \bar{x} splňuje (9.9) pro každé $d \in \mathcal{N}(B)$. Poslední podmínka je ekvivalentní $A\bar{x} - b \in \mathcal{H}(B^T)$, takže existuje $\bar{\lambda} \in \mathbb{R}^m$ tak, že platí (9.10).

Necht $A|\mathcal{N}(B)$ je pozitivně semidefinitní. Jestliže existuje $\bar{\lambda}$ a $\bar{x} \in \Omega$ tak, že platí (9.10), potom podle Taylorova rozvoje (9.8)

$$f(\bar{x} + d) - f(\bar{x}) = L_0(\bar{x} + d, \bar{\lambda}) - L_0(\bar{x}, \bar{\lambda}) = \frac{1}{2}d^T Ad \geq 0$$

pro každé $d \in \mathcal{N}(B)$, takže \bar{x} je řešením (9.3). □

Z věty 9.1 vyplývá, že řešení úlohy (9.3) je řešením soustavy rovnic, kterou si můžeme zapsat v blokovém tvaru

$$\begin{bmatrix} A & B^T \\ B & O \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} b \\ c \end{bmatrix}. \quad (9.12)$$

Dvojici (x, λ) která splňuje (9.12), budeme nazývat *Karush-Kuhn-Tuckerova dvojice* nebo stručně *KKT dvojice*. Soustavu (9.12) nazýváme *Karush-Kuhn-Tuckerovou podmínkou* nebo stručně *KKT podmínkou*.

9.4 Nutné podmínky pro úlohy nelineárního programování

Vraťme se nyní k minimalizaci nelineární funkce s omezením ve tvaru rovností (9.1). Hlavní komplikace je v nelineárních rovnostech, které komplikují charakteristiku přípustných přírůstků argumentu v okolí lokálního řešení \bar{x} , tedy takových vektorů $d \in \mathbb{R}^n$, pro které platí

$$\bar{x} + d \in \Omega$$

Nemusí totiž existovat *přípustné směry*, tj. vektory d pro které platí

$$\bar{x} + \alpha d \in \Omega \quad \text{pro } \alpha \text{ dostatečně malé.}$$

Lze si to snadno ověřit na množině $\Omega = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 2\}$ v $x_1 = x_2 = 1$.

Problematiku si vysvětlíme podrobněji na tečných směrech, které si zavedeme pomocí přípustné křivky. *Křivkou* zde nazýváme množinu

$$\mathcal{C} = \{x(t) : t \in \langle t_0, t_1 \rangle\},$$

kde $t_0 < t_1$ a $x : \mathbb{R} \rightarrow \mathbb{R}^n$ je prosté zobrazení, které má alespoň dvě spojité derivace. Křivka \mathcal{C} se nazývá *přípustnou křivkou* množiny Ω , jestliže $\mathcal{C} \subseteq \Omega$. V dalším se omezíme na křivky, které prochází daným lokálním řešením \bar{x} , takže existuje $\bar{t} \in \langle t_0, t_1 \rangle$ tak, že platí $\bar{x} = x(\bar{t})$. Budeme navíc předpokládat, že taková křivka existuje a $\bar{t} = 0$ (toho můžeme dosáhnout posunutím argumentu t), takže

$$\bar{x} = x(0).$$

První derivace $x(t)$ je vektor

$$x'(t) = \frac{dx}{dt},$$

který určuje směr tečny k $x(t)$ v t . Je-li tedy \bar{x} řešením (9.1), potom

$$f(\bar{x}) = f(x(0)) = \min_{t \in \langle t_0, t_1 \rangle} f(x(t)),$$

takže platí

$$\frac{d}{dt} f(x(t))|_{t=0} = x'(0)^T \nabla f(x(0))|_{t=0} = x'(0)^T \nabla f(\bar{x}) = 0. \quad (9.13)$$

Definice 9.2. Označme si

$$T_{\Omega}(\bar{x}) = \{d \in \mathbb{R}^n : d = x'(0) \text{ pro přípustnou křivku } x \text{ množiny } \Omega\}.$$

Množina $T_{\Omega}(\bar{x})$ doplněná nulovým vektorem se nazývá *tečný kužel*^a v \bar{x} k množině Ω .

^aMnožinu $\mathcal{K} \subseteq \mathbb{R}^n$ nazveme (konvexním) *kuželem*, jestliže $\alpha x \in \mathcal{K}$ pro všechny $\alpha \geq 0$ a $x \in \mathcal{K}$.

Pomocí nového označení a (9.13) můžeme definovat podmínku pro řešení \bar{x} ve tvaru

$$d^T \nabla f(\bar{x}) = 0$$

pro všechna $d \in T_{\Omega}(\bar{x})$, tj.

$$\nabla f(\bar{x}) \perp T_{\Omega}(\bar{x}). \quad (9.14)$$

Praktické použití kužele $T_{\Omega}(\bar{x})$ k charakteristice minima je velmi problematické, neboť jeho popis je komplikovaný a tečný kužel je v obecném případě pouze podmnožinou nulového prostoru matice $\nabla h(\bar{x})^T$. Abychom si ověřili druhé tvrzení, stačí si všimnout, že $h_i(x(t)) = 0$ pro $t \in \langle t_0, t_1 \rangle$, takže pro libovolnou přípustnou křivku $x(t)$ platí

$$h'_i(x(t)) = x'(t)^T \nabla h_i(x(t)) = 0.$$

Použijeme-li tento vztah pro $t = 0$, dostaneme snadno

$$d^T \nabla h_i(\bar{x}) = 0$$

pro všechna $d \in T_\Omega(\bar{x})$, tedy

$$T_\Omega(\bar{x}) \subseteq \mathcal{N}(\nabla h(\bar{x})^T). \quad (9.15)$$

Naštěstí je tečný kužel v typických případech shodný s nulovým prostorem $\nabla h(\bar{x})^T$. Toto tvrzení si snadno ověříme pro lineární vazby definované pomocí $h(x) = Bx - c$. K tomu si stačí všimnout, že $B = \nabla h(x)^T$, takže d patří do jádra Jacobiánu h v libovolném $x \in \mathbb{R}^n$, právě když $Bd = 0$, a pro $x(t) = td$ platí

$$x'(t) = x'(0) = d,$$

tedy $d \in T_\Omega(\bar{x})$ a

$$T_\Omega(\bar{x}) = \mathcal{N}(\nabla h(\bar{x})^T). \quad (9.16)$$

Zobecníme-li si definici Lagrangiánu pomocí

$$L_0(x, \lambda) = f(x) + h(x)^T \lambda,$$

můžeme shrnout naše pozorování do následující věty.

Věta 9.3. *Nechť je problém (9.1) definován pomocí diferencovatelných zobrazení f a h . Nechť problém (9.1) má lokální minimum v \bar{x} a nechť*

$$\mathcal{N}(\nabla h(\bar{x})^T) = T_\Omega(\bar{x}). \quad (9.17)$$

Potom existuje $\bar{\lambda} \in \mathbb{R}^m$ tak, že

$$\nabla_x L_0(\bar{x}, \bar{\lambda}) = \nabla f(\bar{x}) + \nabla h(\bar{x})^T \bar{\lambda} = 0. \quad (9.18)$$

Podmínka (9.17) je splněna, když h je afinní zobrazení.

Důkaz. Z předpokladu (9.14) a (9.17) plyne, že $\nabla f(\bar{x})$ je ortogonální k $\mathcal{N}(\nabla h(\bar{x})^T)$, tj. $\nabla f(\bar{x}) \in \mathcal{H}(\nabla h(\bar{x}))$. \square

Pomocí věty o implicitní funkci a důvtipu lze dokázat rovnost (9.17) i pro případ, že hodnota $\nabla h(\bar{x})$ je rovna m . My si toto poslední tvrzení odvodíme jiným způsobem v 11. kapitole.

Dvojici $(\bar{x}, \bar{\lambda})$, která splňuje (9.18), budeme opět nazývat *Karush-Kuhn-Tuckerova dvojice* nebo stručně *KKT dvojice*. Soustavu (9.18) s podmínkou $h(x) = 0$ nazýváme *Karush-Kuhn-Tuckerovou podmínkou* nebo stručně *KKT podmínkou*.

Podmínku optimality pro úlohu s rovnostním omezením lze zavést i pomocí jiné geometrické konstrukce než té, která využívá tečný kužel. K tomu si musíme zavést další charakteristiku omezujících podmínek a tou je normálový kužel.

Definice 9.4. Množinu $N_\Omega(\bar{x})$ nazveme *normálový kužel* v \bar{x} k množině Ω , pokud je ortogonálním doplňkem tečného kuželu v \bar{x} k množině Ω , tzn.

$$N_\Omega(\bar{x}) = \{v : v^T w \leq 0 \quad \forall w \in T_\Omega(\bar{x})\}. \quad (9.19)$$

Pomocí normálového kužele můžeme podmínku optimality shrnout v následující větě.

Věta 9.5. *Nechť je problém (9.1) definován pomocí diferencovatelných zobrazení f a h . Nechť problém (9.1) má lokální minimum v \bar{x} , pak*

$$0 \in \nabla f(\bar{x}) + N_\Omega(\bar{x}). \quad (9.20)$$

Poznámka 9.6. Pokud budeme vyžadovat, aby byla v bodě \bar{x} splněna přísnější podmínka pro rovnostní omezení h , dostaneme dokonce nejen existenci (tato podmínka totiž zaručuje platnost podmínky (9.17)), ale navíc i jednoznačnost λ splňující podmínku (9.18). Touto podmínkou je regularita bodu \bar{x} . Pro rovnostní vazby ji zavědeme v následující definici.

Definice 9.7. Mějme minimalizační úlohu (9.1). Přípustný bod x nazveme *regularní*, jestliže gradienty rovnostních vazeb jsou lineárně nezávislé. Tzn. soustava

$$\alpha_1 \nabla h_1(x) + \alpha_2 \nabla h_2(x) + \dots + \alpha_m \nabla h_m(x) = 0$$

má jediné řešení

$$\alpha_1 = \alpha_2 = \dots = \alpha_m = 0.$$

Poznámka 9.8. Přípustný bod, který splňuje podmínku regularity, bývá v některých textech označován jako bod splňující klasifikaci omezení na lineární nezávislost¹.

Poznámka 9.9. Je dobré si uvědomit, že podmínka regularity v přípustném bodě x nám zaručuje jednoznačnost λ splňující podmínku (9.18), ale v případě „větších“ úloh je značně komplikované ji ověřit, a proto je v tomto případě její použití omezené.

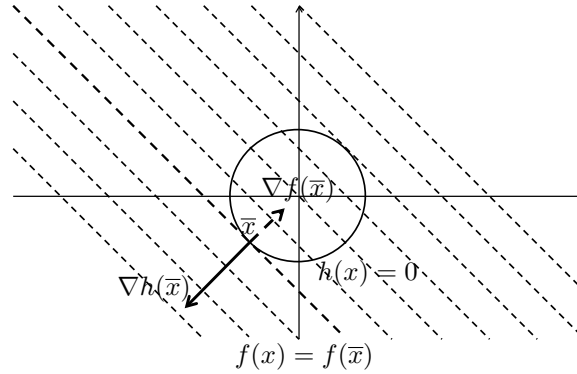
Příklad 9.10. Mějme následující minimalizační úlohu s jedním rovnostním omezením

$$\min_{(x_1, x_2) \in \Omega} x_1 + x_2, \quad \Omega = \{x \in \mathbb{R} : x_1^2 + x_2^2 = 2\}.$$

Napište Karush-Kuhn-Tuckerovu podmínku pro tuto úlohu a najděte Karush-Kuhn-Tuckerovy dvojice, které ji splňují.

¹V anglické literatuře najde čtenář tuto podmínku pod názvem Linear Independence Constraint Qualification či pod zkratkou LICQ.

Řešení. Pomůžeme si obrázkem 9.3, na kterém jsou znázorněny čárkovaně vrstevnice funkce f a nepřerušovanou čarou je znázorněna kružnice, která představuje přípustnou množinu. Minimum je dosaženo v bodě \bar{x} .



Obr. 9.3: Minimum lineární funkce na přípustné množině

Karush-Kuhn-Tuckerova podmínka pro tuto úlohu je ve tvaru

$$\nabla_{(x_1, x_2)} (x_1 + x_2) + \lambda \nabla_{(x_1, x_2)} (x_1^2 + x_2^2 - 2) = 0, \quad x_1^2 + x_2^2 = 2.$$

Z tohoto systému dostaneme po derivování následující systém rovnic

$$1 + 2\lambda x_1 = 0, \quad 1 + 2\lambda x_2 = 0, \quad x_1^2 + x_2^2 - 2 = 0.$$

Jejich řešením dostaneme složky řešení

$$\bar{x}_1 = \bar{x}_2 = -1, \quad \lambda = 0,5 \quad \text{a} \quad \bar{x}_1 = \bar{x}_2 = 1, \quad \lambda = -0,5.$$

Dá se ukázat, že řešením minimalizační úlohy je bod o složkách

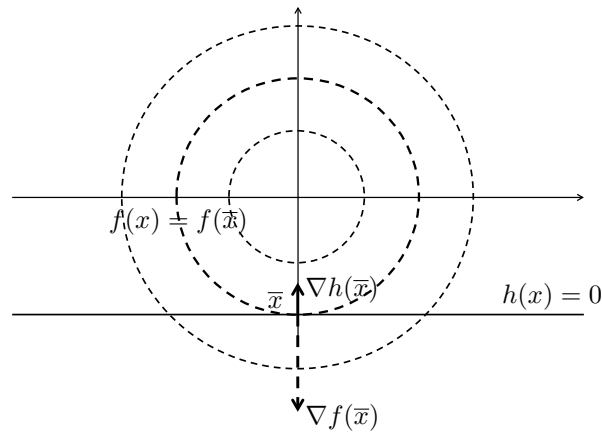
$$\bar{x}_1 = \bar{x}_2 = -1.$$

▲

Příklad 9.11. Mějme následující minimalizační úlohu s jedním rovnostním omezením

$$\min_{(x_1, x_2) \in \Omega} x_1^2 + x_2^2, \quad \Omega = \{x \in \mathbb{R} : x_2 = -5\}.$$

Napište Karush-Kuhn-Tuckerovu podmínku pro tuto úlohu a najděte Karush-Kuhn-Tuckerovy dvojice, které ji splňují.



Obr. 9.4: Minimum kvadratické funkce na přípustné množině

Řešení. Pomůžeme si obrázkem 9.4, na kterém jsou znázorněny čárkované kruhové vrstevnice funkce f a nepřerušovanou čarou je znázorněna přímka, která představuje přípustnou množinu. Minimum je dosaženo v bodě \bar{x} .

Karush-Kuhn-Tuckerova podmínka pro tuto úlohu je ve tvaru

$$\nabla_{(x_1, x_2)} (x_1^2 + x_2^2) + \lambda \nabla_{(x_1, x_2)} (x_2 + 5) = 0, \quad x_2 = -5.$$

Z tohoto systému dostaneme po derivování následující systém rovnic

$$2x_1 = 0, \quad 2x_2 + \lambda = 0, \quad x_2 + 5 = 0.$$

Jejich řešením dostaneme složky řešení

$$\bar{x}_1 = 0, \quad \bar{x}_2 = -5, \quad \lambda = 10.$$

Dá se ukázat, že řešením minimalizační úlohy je bod o složkách

$$\bar{x}_1 = 0, \quad \bar{x}_2 = -5.$$

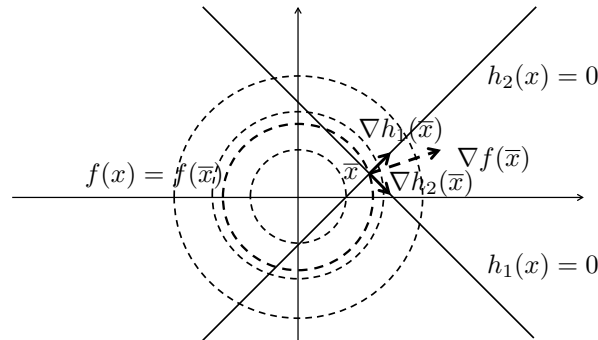
▲

Příklad 9.12. Mějme následující minimalizační úlohu se dvěma rovnostními omezeními

$$\min_{(x_1, x_2) \in \Omega} x_1^2 + x_2^2, \quad \Omega = \{x \in \mathbb{R} : x_1 + x_2 = 2, \quad x_1 - x_2 = 1\}.$$

Napište Karush-Kuhn-Tuckerovu podmínku pro tuto úlohu a najděte Karush-Kuhn-Tuckerovy dvojice, které ji splňují.

Řešení. Pomůžeme si obrázkem 9.5, na kterém jsou znázorněny čárkované kruhové vrstevnice funkce f a nepřerušovanou čarou jsou znázorněny přímky, jejichž průsečík představuje přípustnou množinu. Minimum je dosaženo v bodě \bar{x} .



Obr. 9.5: Minimum kvadratické funkce na přípustné množině

Karush-Kuhn-Tuckerova podmínka pro tuto úlohu je ve tvaru

$$\begin{aligned} \nabla_{(x_1, x_2)} (x_1^2 + x_2^2) + \lambda_1 \nabla_{(x_1, x_2)} (x_1 + x_2 - 2) + \lambda_2 \nabla_{(x_1, x_2)} (x_1 - x_2 - 1) &= 0, \\ x_1 + x_2 &= 2, \quad x_1 - x_2 = 1. \end{aligned}$$

Z tohoto systému dostaneme po derivování následující systém rovnic

$$2x_1 + \lambda_1 + \lambda_2 = 0, \quad 2x_2 + \lambda_1 - \lambda_2 = 0, \quad x_1 + x_2 = 2, \quad x_1 - x_2 = 1.$$

Jejich řešením dostaneme složky řešení

$$\bar{x}_1 = 1,5, \quad \bar{x}_2 = 0,5, \quad \lambda_1 = -2, \quad \lambda_2 = -1.$$

Dá se ukázat, že řešením minimalizační úlohy je bod o složkách

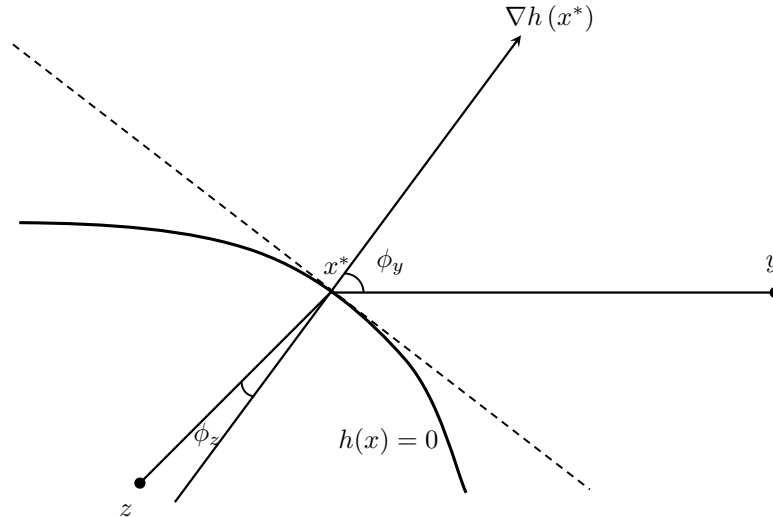
$$\bar{x}_1 = 1,5, \quad \bar{x}_2 = 0,5.$$

▲

9.5 Odvození Snellova zákona lomu

Nyní si ukážeme aplikaci Karush-Kuhn-Tuckerových podmínek při odvození Snellova zákona lomu. Uvažujme hladkou dvojrozměrnou křivku popsanou rovnicí $h(x) = 0$, kde funkce $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ je spojitě diferencovatelná. Nechť tato křivka rozděljuje jako na obrázku 9.6 rovinu na dvě oblasti a rychlost světla je v každé z těchto oblastí různá. Dále nechť y a z jsou body ležící na opačných stranách křivky $h(x) = 0$. Předpokládejme, že světlo, které projde cestu z bodu y do z , protíná křivku $h(x) = 0$ v bodě x^* . Označme rychlost světla v oblasti s bodem y jako v_y a rychlost světla v oblasti s bodem z jako v_z , tzn. že světelný paprsek projde úsečku $\overrightarrow{yx^*}$ rychlostí v_y

a úsečku $\overrightarrow{x^*z}$ rychlostí v_z . Normála ke křivce $h(x) = 0$ v bodě x^* je popsána vektorem $\nabla h(x^*)$. Dále si označme úhel mezi normálou ke křivce $h(x) = 0$ v bodě x^* a úsečkou $\overrightarrow{yx^*}$ jako ϕ_y a úhel mezi normálou ke křivce $h(x) = 0$ v bodě x^* a úsečkou $\overrightarrow{zx^*}$ jako ϕ_z . Vše je ilustrováno opět na obrázku 9.6.



Obr. 9.6: Snellův zákon lomu

Dle Fermatova principu projde světlo z bodu y do z po cestě s minimální dobou průchodu světla. Na základě toho můžeme najít polohu bodu x^* ležícího na křivce $h(x) = 0$ jako řešení minimalizační úlohy s omezením na rovnost

$$\min_{h(x)=0} f(x), \quad (9.21)$$

kde funkce f popisuje čas, za který světelný paprsek projde z bodu y přes bod x do bodu z . Funkce f je dána takto

$$f(x) = \frac{\|y - x\|}{v_y} + \frac{\|z - x\|}{v_z}.$$

Pro nalezení bodu x^* , který je řešením úlohy (9.21), použijeme větu 9.3 a podmínku (9.18). Vzhledem k tomu, že předpokládáme $\nabla h(x^*) \neq 0$, je bod x^* regulární a z toho plyne existence x^* a λ^* , pro které platí (dle věty 9.3 a podmínky (9.18)) Karush-Kuhn-Tuckerova podmínka

$$\nabla_x L_0(x^*, \lambda^*) = \nabla f(x^*) + \nabla h(x^*)^T \lambda^* = 0. \quad (9.22)$$

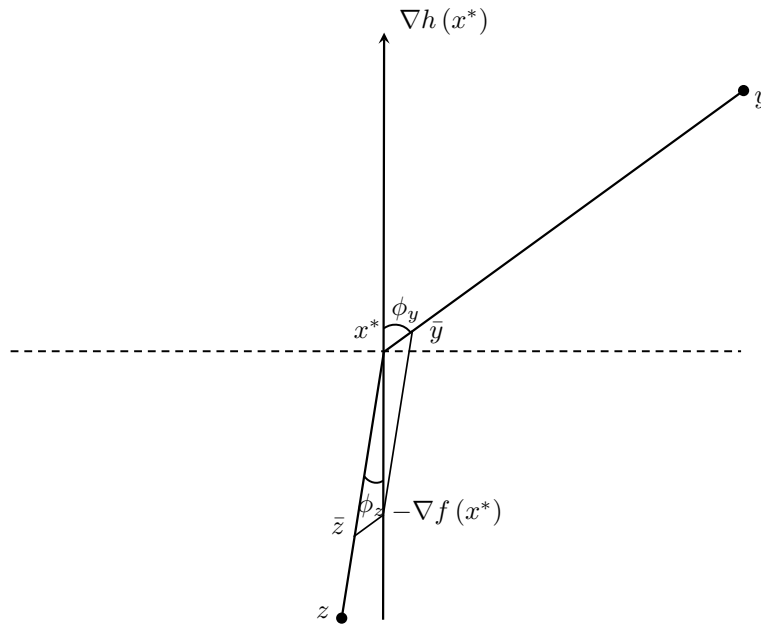
To znamená, že vektor $\nabla f(x^*)$ je skalární násobek vektoru $\nabla h(x^*)$. Vektor

$$\nabla f(x^*) = \frac{x^* - y}{v_y \|x^* - y\|} + \frac{x^* - z}{v_z \|x^* - z\|}. \quad (9.23)$$

Abychom mohli použít podmínku (9.22), zavedeme vektory

$$\bar{y} = x^* + \frac{y - x^*}{v_y \|y - x^*\|} \quad \bar{z} = x^* + \frac{z - x^*}{v_z \|z - x^*\|}. \quad (9.24)$$

Navíc zavedeme rovnoběžník se stranami $\overrightarrow{x^* \bar{y}}$ a $\overrightarrow{x^* \bar{z}}$. Podle (9.23) je vektor $-\nabla f(x^*)$ úhlopříčkou tohoto rovnoběžníku a navíc je podle (9.22) tento vektor kolmý ke křivce $h(x) = 0$ v bodě x^* . Vše je zobrazeno na obrázku 9.7.



Obr. 9.7: Snellův zákon lomu - odvození

Z rovnoběžnosti stran rovnoběžníku plyne, že vzdálenosti bodů \bar{y} a \bar{z} od přímky, na které leží vektory $-\nabla f(x^*)$ a $\nabla h(x^*)$, jsou stejné. Označme tyto vzdálenosti jako l . Z obrázku 9.7 plyne

$$\sin \phi_y = \frac{l}{\|\bar{y} - x^*\|}, \quad \sin \phi_z = \frac{l}{\|\bar{z} - x^*\|}.$$

Protože $\|\bar{y} - x^*\| = \frac{1}{v_y}$ a $\|\bar{z} - x^*\| = \frac{1}{v_z}$ platí

$$l = \sin \phi_y \|\bar{y} - x^*\| = \frac{\sin \phi_y}{v_y}, \quad l = \sin \phi_z \|\bar{z} - x^*\| = \frac{\sin \phi_z}{v_z}.$$

Z toho již snadno odvodíme

$$\frac{\sin \phi_y}{v_y} = \frac{\sin \phi_z}{v_z},$$

což je hledaný *Snellův zákon lomu*.

Kapitola 10

Metody řešení úloh s rovnostmi

V této kapitole se seznámíme se základními algoritmy pro nalezení minima nelineární funkce vzhledem k omezením ve tvaru rovností. Seznámíme se postupně se třemi metodami redukce naší úlohy na úlohu minimalizace bez omezení, s modifikací gradientní metody pro řešení úloh s omezením. V klasické analýze se řešení úloh minimalizace s rovnostními vazbami nazývají také *podmíněné extrémy*. Pro jednoduchost budeme předpokládat, že f je konvexní a že úloha má jediné řešení.

10.1 Vnější penalta

Pravděpodobně nejjednodušší metoda převedení úlohy s omezením (9.1) spočívá v zahrnutí rovnostních vazeb do cenové funkce tak, že k ní přičteme vhodný člen, který penalizuje porušení vazeb. Omezíme se zde na nejčastěji používanou metodu *kvadratické penalty*, která aproximuje řešení \bar{x} úlohy (9.1) řešením \bar{x}_ρ úlohy

$$\min_{x \in \mathbb{R}^n} f_\rho(x), \quad f_\rho(x) = f(x) + \frac{\rho}{2} \|h(x)\|^2, \quad (10.1)$$

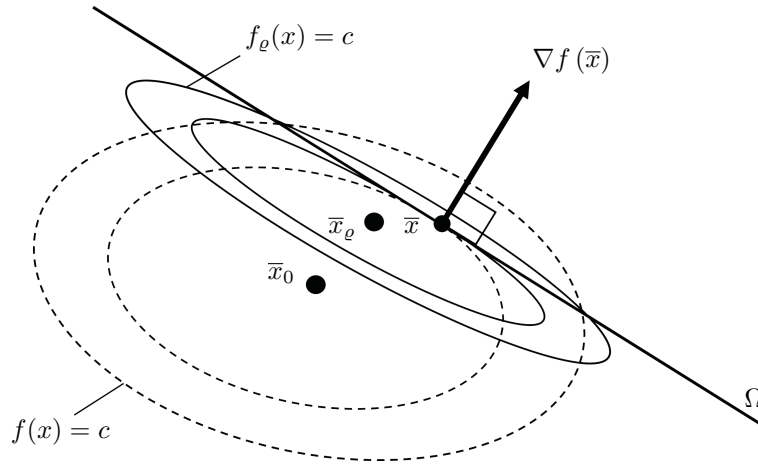
kde $\rho > 0$ je penalizační parametr a $\|h(x)\|^2$ je penalizační funkce.

Intuitivně je zřejmé, že když je penalizační parametr ρ velký, tak řešení \bar{x}_ρ , ve kterém je dosaženo minima penalizované funkce f_ρ , nemůže být daleko od přípustné množiny. Je dokonce zřejmé, že kdyby $\rho = \infty$, pak by bylo minimum f_ρ řešením původní úlohy (9.1). Můžeme tedy očekávat, že pro dostatečně velké hodnoty penalizačního parametru ρ bude řešení \bar{x} úlohy (9.1) blízko \bar{x}_ρ . Efekt penalizace je na obr. 10.1. Z obrázku je patrné, že řešení penalizované úlohy je typicky blízko přípustné množiny, avšak nepatří do ní. Proto se naše penalizační metoda nazývá také *metoda externí penalty*.

Penalizační efekt si ukážeme na *rozšířeném Lagrangiánu*

$$L_\rho(x, \lambda) = L_0(x, \lambda) + \frac{\rho}{2} \|h(x)\|^2 = f(x) + h(x)^T \lambda + \frac{\rho}{2} \|h(x)\|^2. \quad (10.2)$$

Pro $x \in \Omega$ tedy platí $L_0(x, \lambda) = L_\rho(x, \lambda)$. Předpokládejme, že $\lambda \in \mathbb{R}^m$ je libovolné,



Obr. 10.1: Význam kvadratické penalty

ale předem zvolené, a že $\min L_0(x, \lambda)$ existuje a je dosaženo v jediném argumentu, takže si můžeme označit

$$\bar{x}_0 = \arg \min_x L_0(x, \lambda) \quad \text{a} \quad \bar{x}_\rho = \arg \min_x L_\rho(x, \lambda).$$

Potom řešení \bar{x} úlohy (9.1) splňuje

$$L_0(\bar{x}_0, \lambda) \leq L_0(\bar{x}_\rho, \lambda) + \frac{\rho}{2} \|h(\bar{x}_\rho)\|^2 = L_\rho(\bar{x}_\rho, \lambda) \leq L_\rho(\bar{x}, \lambda) = f(\bar{x}),$$

takže s použitím $L_0(\bar{x}_0, \lambda) \leq L_0(\bar{x}_\rho, \lambda)$ dostaneme

$$\|h(\bar{x}_\rho)\|^2 \leq \frac{2}{\rho} (f(\bar{x}) - L_0(\bar{x}_\rho, \lambda)) \leq \frac{2}{\rho} (f(\bar{x}) - L_0(\bar{x}_0, \lambda)). \quad (10.3)$$

Odtud snadno plyne, že narušení rovnostních vazeb $\|h(\bar{x}_\rho)\|$ může být libovolně malé při dostatečně velkém ρ .

Než se podíváme, jak \bar{x}_ρ splňuje nutnou podmínku minima (9.18), připomeňme si, že rozšířený Lagrangian je definován pomocí (10.2), takže

$$\nabla_x L_\rho(x, \lambda) = \nabla f(x) + \nabla h(x)\lambda + \rho \nabla h(x)h(x) = \nabla f(x) + \nabla h(x)(\lambda + \rho h(x))$$

a \bar{x}_ρ splňuje

$$\nabla_x L_\rho(\bar{x}_\rho, \lambda) = \nabla f(\bar{x}_\rho) + \nabla h(\bar{x}_\rho)(\lambda + \rho h(\bar{x}_\rho)) = o.$$

Zavedeme-li si označení

$$\tilde{\lambda} = \lambda + \rho h(x), \quad (10.4)$$

potom můžeme poslední rovnost přepsat ve tvar

$$\nabla f(\bar{x}_\rho) + \nabla h(\bar{x}_\rho)\tilde{\lambda} = o. \quad (10.5)$$

takže $(\bar{x}_\rho, \tilde{\lambda})$ splňuje nutnou podmínku minima (9.18) přesně. Je-li λ nějaká aproximace vektoru Lagrangeových multiplikátorů řešení úlohy (9.1), potom z našich úvah vyplývá, že $\tilde{\lambda}$ je lepší aproximace vektoru Lagrangeových multiplikátorů řešení úlohy (9.1). Z nerovnosti (10.3) lze odvodit, že $(\bar{x}_\rho, \tilde{\lambda})$ může za rozumných předpokladů aproximovat řešení úlohy (9.1) i příslušný vektor Lagrangeových multiplikátorů s libovolnou přesností.

10.2 Metoda oboru hodnot pro kvadratické programování

Jak jsme si ukázali v sekci 9.3, KKT dvojice řešení úlohy (9.3) splňuje

$$\begin{bmatrix} A & B^T \\ B & O \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} b \\ c \end{bmatrix}.$$

Je-li matice A pozitivně definitní, můžeme první blokový řádek přenásobit $-BA^{-1}$ a pak přičíst výsledek ke druhému řádku, čímž vyloučíme x , a po změně znaménka dostaneme soustavu

$$BA^{-1}B^T\lambda = BA^{-1}b - c,$$

která neobsahuje x . Jejím řešením dostaneme multiplikátor řešení $\bar{\lambda}$, který je jediný, právě když jsou řádky B lineárně nezávislé. Po dosazení $\bar{\lambda}$ do první rovnice dostaneme řešení \bar{x} ve tvaru

$$\bar{x} = A^{-1}(b - B^T\bar{\lambda}).$$

Pokud je dimenze úlohy velká a pokud má matice A speciální strukturu, nemusí být výhodné sestavovat matici $BA^{-1}B^T$ explicitně. Alternativou je řešení pomocí vhodné iterační metody, na příklad pomocí metody sdružených gradientů s předpoklady. Tento postup se často používá při řešení velkých (řádově miliony až miliardy) úloh kvadratického programování vznikajících při řešení úloh mechaniky či magnetostatiky pomocí metod rozložení oblasti.

10.3 Metoda nulového prostoru pro kvadratické programování

Alternativou k metodě oboru hodnot (range space method) je metoda nulového prostoru (null space method), kterou můžeme použít, když máme bázi Z nulového prostoru $\mathcal{N}(B)$ a x_0 takové, že platí

$$Bx_0 = c.$$

Jelikož $\Omega = \{x_0 + Zy : y \in \mathbb{R}^d\}$, kde d je dimenze nulového prostoru $\mathcal{N}(B)$, můžeme dosadit do (9.3) a zredukovat tak původní problém na problém

$$\min_{x \in \Omega} f(x) = \min_{y \in \mathbb{R}^d} f(x_0 + Zy) = \frac{1}{2}y^T Z^T A Z y - (b - Ax_0)^T Z y + \frac{1}{2}x_0^T A x_0 - b^T x_0.$$

Řešení tohoto problému bez omezení splňuje gradientní podmínku

$$Z^T A Z y = Z^T (b - Ax_0), \quad (10.6)$$

jejíž řešení vypočteme jako

$$\bar{y} = (Z^T A Z)^{-1} Z^T (b - Ax_0),$$

a řešení problému (9.3) získáme pomocí

$$\bar{x} = x_0 + Z\bar{y}.$$

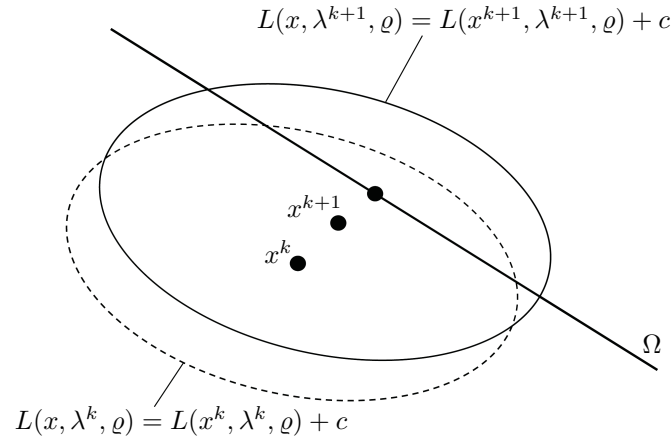
Pokud řešíme soustavu (10.6) metodou sdružených gradientů, potom můžeme metodu nulového prostoru použít beze změny i pro řešení úlohy s pozitivně semidefinitní maticí A a B se závislými řádky.

10.4 Metody Uzawova typu pro kvadratické programování

Metoda penalty se sice často používá pro svoji jednoduchost, avšak získání dobré aproximace řešení touto metodou vyžaduje velké hodnoty penalizačního parametru ϱ , což často vede k vážným problémům při počítačové implementaci. Možnou nápravu si vysvětlíme na následující analýze řešení úlohy kvadratického programování s pozitivně definitním Hessiánem.

Předpokládejme, že máme řešení x_ϱ penalizovaného problému (10.1). Potom můžeme modifikovat f pomocí vhodného λ tak, že minimum bez omezení modifikované cenové funkce $\bar{f} = L_0$ bez penalizačního členu je dosaženo znovu v x_ϱ . Pak můžeme čekat, že najdeme lepší aproximaci řešení tak, že přičteme penalizační člen k \bar{f} , čímž dostaneme L_ϱ , třeba se stejnou hodnotou penalizačního parametru, a budeme hledat

argument minima \bar{f}_ϱ tak, jak je naznačeno na obr. 10.2. Výsledkem je klasická *metoda rozšířených (augmented) Lagrangianů*, známá též jako *metoda multiplikátorů*, která byla navržena Hestenesem a Powelem v roce 1969.



Obr. 10.2: Iterace metody rozšířených Lagrangianů

Popíšeme si zde variantu metody rozšířených Lagrangianů jejímž speciálními případy jsou *Uzawiv algoritmus* z roku 1958, který dostaneme volbou $\varrho_k = 0$, $k = 0, 1, \dots$, a původní *metoda multiplikátorů*, která odpovídá $r_k = \varrho_k$. Náš algoritmus metody rozšířených Lagrangianů si můžeme naznačit v následujícím tvaru.

Algoritmus 10.1 Metoda rozšířených Lagrangianů.

Mějme symetrickou pozitivně definitní matici $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ a $c \in \mathcal{H}(B)$.

Krok 0. {Inicializace.}

Zvolme $\lambda^0 \in \mathbb{R}^m$, $r > 0$, $r_k \geq r$, $\varrho_k \geq 0$

for $k = 0, 1, 2, \dots$

Krok 1. {Minimalizace vzhledem k proměnné x .}

$x^k = \arg \min\{L(x, \lambda^k) : x \in \mathbb{R}^n\}$

Krok 2. {Aktualizace Lagrangeových multiplikátorů.}

$\lambda^{k+1} = \lambda^k + r_k(Bx^k - c)$

end for

Jelikož v x^k je dosaženo minimum Lagrangianu L vzhledem k první proměnné, platí

$$\nabla_x L(x^k, \lambda^k) = (A + \varrho_k B^T B)x^k - b - \varrho_k B^T c + B^T \lambda^k = 0,$$

takže krok 1 algoritmu 10.1 může být implementován řešením soustavy

$$(A + \varrho_k B^T B)x^k = b + \varrho_k B^T c - B^T \lambda^k. \quad (10.7)$$

Abychom pochopili nový algoritmus lépe, podíváme se na jeho jinou formulaci, kterou dostaneme eliminací x^k z algoritmu 10.1. Označíme-li si pro každé $\varrho \in \mathbb{R}$

$$A_\varrho = A + \varrho B^T B,$$

můžeme použít (10.7) k odvození

$$x^k = A_{\varrho_k}^{-1}(b + \varrho_k B^T c - B^T \lambda^k).$$

Po substituci x^k do kroku 2 algoritmu 10.1 a jednoduchých manipulací dostaneme náš algoritmus rozšířených Lagrangiánů ve tvaru:

$$\begin{aligned} & \text{Zvolme } \lambda^0 \in \mathbb{R}^m, \\ & \lambda^{k+1} = \lambda^k - r_k (BA_{\varrho_k}^{-1} B^T \lambda^k - BA_{\varrho_k}^{-1} (b + \varrho_k B^T c) + c). \end{aligned} \quad (10.8)$$

Výraz ve vnější závorce (10.8) je gradient funkce

$$\theta_\varrho(\lambda) = \frac{1}{2} \lambda^T BA_\varrho^{-1} B^T \lambda - (BA_\varrho^{-1} (b + \varrho B^T c) - c) \lambda,$$

neboť platí

$$\nabla \theta_{\varrho_k}(\lambda^k) = BA_{\varrho_k}^{-1} B^T \lambda^k - BA_{\varrho_k}^{-1} (b + \varrho_k B^T c) + c. \quad (10.9)$$

Porovnáme-li poslední výraz s (10.8), můžeme si druhý krok Uzawova algoritmu přepsat na tvar

$$\lambda^{k+1} = \lambda^k - r_k \nabla \theta_{\varrho_k}(\lambda^k).$$

Iterace algoritmu metody rozšířených Lagrangiánů je tedy totožná s krokem gradientní metody pro hledání $\min \theta_\varrho(\lambda)$ s délkou kroku r_k . Poznamenejme zde, že funkce θ_ϱ je až na znaménko a konstantu totožná s duální funkcí spojenou s penalizovaným problémem (10.1) a platí

$$\lambda^k = \arg \min_{\lambda} \theta_{\varrho_k}(\lambda).$$

Opakováním úvah ze sekce 4.1 dostaneme

$$\|\lambda^{k+1} - \bar{\lambda}\| \leq \max\{1 - r^k \lambda_{\min}(BAB^T), r_k \lambda_{\max}(BAB^T) - 1\} \|\lambda^k - \bar{\lambda}\|,$$

kde $\lambda_{\min}(BAB^T)$ a $\lambda_{\max}(BAB^T)$ značí nejmenší a největší vlastní čísla matice BAB^T . Další výsledky lze získat analýzou spektra matice BAB^T .

10.5 SMALE

Tuto kapitolu ukončíme popisem algoritmu SMALE, který připouští přibližné řešení minimalizační úlohy v prvním kroku algoritmu rozšířených Lagrangiánů. Nepřesný algoritmus je založen na pozorování, že pro dostatečně velké r a *přibližné* minimum x_ϱ rozšířeného Lagrangiánu $L_\varrho(x, \lambda)$ vzhledem k x můžeme modifikovat λ tak, že x_ϱ je také *přibližné* minimum $L_0(x, \tilde{\lambda})$, kde $\tilde{\lambda} = \lambda + r(Bx - c)$. Tak můžeme doufat, že najdeme lepší aproximaci multiplikátoru minimalizací $L_\varrho(x, \tilde{\lambda})$. Jelikož lepší aproximaci odpovídá větší hodnota Lagrangiánu, je přirozené zvyšovat hodnotu penalizačního parametru (nebo zvyšovat přesnost řešení vnitřní úlohy), dokud nebude zajištěn růst Lagrangiánu v následujících iteracích. Teorie ukazuje, že hodnota penalizačního parametru ϱ (nebo parametru M) není velká (malá) a nezávisí na matici B . Celý algoritmus pak vypadá následovně.

Algoritmus 10.2 Metoda semimonotónních rozšířených Lagrangiánů - Semimonotonic augmented Lagrangians (SMALE).

Dána SPD matice $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^n$ a $c \in \mathcal{H}(B)$.

Krok 0. {Inicializace.}

Zvolte $\eta > 0$, $\beta > 1$, $M > 0$, $\varrho_0 > 0$, $\lambda^0 \in \mathbb{R}^m$

for $k = 0, 1, 2, \dots$

Krok 1. {Vnitřní iterace s adaptivní kontrolou přesnosti.}

Nalezněte x^k tak, aby platilo

$$\|g(x^k, \lambda^k, \varrho_k)\| \leq \min\{M\|Bx^k - c\|, \eta\}.$$

Krok 2. {Aktualizace Lagrangeových multiplikátorů.}

$$\lambda^{k+1} = \lambda^k + \varrho_k(Bx^k - c)$$

Krok 3. {Aktualizace ϱ v případě že růst Lagrangiánu není dostatečný.}

if $k > 0$ a

$$L(x^k, \lambda^k, \varrho_k) < L(x^{k-1}, \lambda^{k-1}, \varrho_{k-1}) + \frac{\varrho_k}{2} \|Bx^k - c\|^2$$

$$\varrho_{k+1} = \beta \varrho_k$$

else

$$\varrho_{k+1} = \varrho_k.$$

end if

end for

V kroku 1 lze použít jakýkoliv konvergentní algoritmus pro minimalizaci ryze konvexní funkce jako je metoda sdružených gradientů s předpokmáním.

Příklady k procvičení

1. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = x_1^4 + x_1^2 + x_2^2.$$

Řešte úlohu optimalizace s omezením na rovnost

$$\min_{x \in \Omega} f(x_1, x_2), \quad \Omega = \{x \in \mathbb{R}^2 : x_1 + 3x_2 - 5 = 0\}.$$

Uvedenou úlohu s omezením na rovnost řešte pomocí metody vnější penalty.

2. Mějme funkci

$$f(x_1, x_2, x_3) = x_1^4 + x_1^2 + x_2^2 + x_3^4.$$

Řešte úlohu optimalizace s omezeními na rovnost

$$\min_{x \in \Omega} f(x_1, x_2, x_3), \quad \Omega = \{x \in \mathbb{R}^3 : x_1 + 3x_2 + 3x_3 - 5 = 0, 2x_1 - x_3 - 3 = 0\}.$$

Uvedenou úlohu s omezeními na rovnost řešte pomocí metody vnější penalty.

3. Minimum kvadratické funkce

$$f(x) = \frac{1}{2}x^T Ax - b^T x$$

aproximuje (až na jednotky) průhyb struny jednotkové délky zatížené jednotkovou silou a uchycené na obou koncích ($x_1 = x_n = 0$), kde $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$ a

$$A = (n-1) \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$b = -\frac{1}{n-1} \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 1 \\ 0 \end{bmatrix}.$$

Najděte průhyb struny v případě, že její druhá čtvrtina je vázaná tak, že se pohybuje vodorovně, tedy

$$x_{[n/4]} = x_{[n/4+1]} = \cdots = x_{[n/2]}.$$

Uvedenou úlohu s omezeními na rovnost řešte pomocí metody vnější penalty.

Kapitola 11

Podmínky minima pro úlohy s omezením ve tvaru nerovností

V této kapitole se budeme zabývat řešením dalšího typu optimalizačních úloh s omezením. V technických úlohách takové omezení vzniká například potřebou zahrnout do formulace úlohy omezení zdrojů, které máme k dispozici, požadavky na pevnost konstrukcí nebo omezení objemu či tvaru hledaného optimálního tělesa při úloze tvarové optimalizace. Nyní budeme pracovat s omezeními ve tvaru nerovností a ukážeme si, jak najít řešení úlohy

$$\min_{x \in \Omega} f(x), \quad \Omega = \{x \in \mathbb{R}^n : g(x) \leq o\}, \quad g : \mathbb{R}^n \rightarrow \mathbb{R}^r, \quad r < n. \quad (11.1)$$

Budeme při tom předpokládat, že f je dvakrát spojitě diferencovatelná funkce a g je spojitě diferencovatelná. Množina Ω se nazývá přípustná množina. Navíc zde zavedeme množinu *aktivních nerovnostních vazeb* v bodě x (nebo také *aktivní množinu* v bodě x). Tuto množinu označíme jako $\mathcal{A}(x)$ a definujeme ji

$$\mathcal{A}(x) := \{i : g_i(x) = 0\}.$$

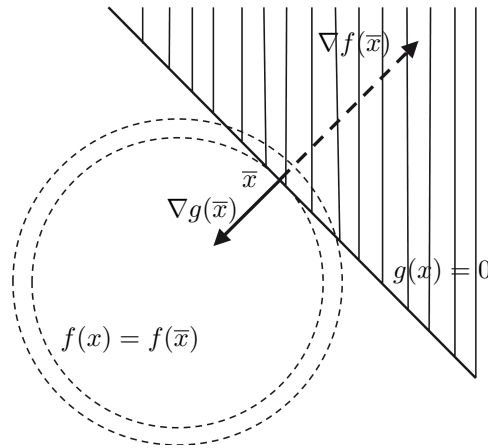
11.1 Geometrická ilustrace podmínky minima

Začneme s řešením jednoduché úlohy

$$\min_{g(x) \leq 0} f(x), \quad f(x) = x_1^2 + x_2^2, \quad g(x) = 4 - x_1 - x_2. \quad (11.2)$$

Pomůžeme si obrázkem 11.1, na kterém jsou znázorněny čárkovaně kruhové vrstevnice funkce f a nepřerušovanou čarou je znázorněna přímka, která představuje hranici přípustné množiny. Přípustná množina je vyšrafována plnými čarami. Minimum je dosaženo v bodě \bar{x} .

Jelikož minimum f bez omezení je dosaženo ve společném středu všech kružnic popsanych pomocí rovnic $f(x) = c$, $c > 0$, je z obrázku vidět, že nejmenší z kružnic,



Obr. 11.1: Minimum kvadratické funkce na polorovině

na nichž leží alespoň jeden přípustný bod, je právě ta kružnice, která se dotýká přípustné množiny. Ve společném bodě \bar{x} platí, že gradient f je násobkem gradientu g , tedy existuje $\mu \in \mathbb{R}$ tak, že

$$\nabla f(\bar{x}) + \mu \nabla g(\bar{x}) = 0.$$

Dostali jsme tak algebraickou formu našeho pozorování. Jelikož řešení musí být přípustné, můžeme napsat rovnici a nerovnici, které v našem případě určují řešení:

$$\nabla f(x) + \mu \nabla g(x) = 0, \quad g(x) \leq 0.$$

Jelikož rovnice je vektorová, můžeme si systém přepsat jako dvě skalární rovnice a jednu nerovnici

$$2x_1 - \mu = 0, \quad 2x_2 - \mu = 0, \quad x_1 + x_2 \geq 4.$$

Jejich řešením dostaneme složky řešení $\bar{x}_1 = \bar{x}_2 = 2$, $\mu = 4$. Povšimněte si, že jsme získali nejen řešení úlohy (11.2), ale také hodnotu μ , na kterou se nás nikdo neptal. V dalším výkladu uvidíme, že μ je důležitá veličina, jejíž pochopení nám umožní zavést nové algoritmy a dodatečnou informaci o kvalitě řešení. Také si všimněme, že μ je pro naši úlohu kladné.

11.2 Úloha kvadratického programování s nerovnostmi a Lagrangián

Předpokládejme, že máme úlohu kvadratického programování s přípustnou množinou, která je určena lineárními nerovnostmi. Takové množiny nazýváme také *polyedrálními množinami*. Tzn. máme úlohu

$$\min_{x \in \Omega} f(x), \quad f(x) = x^T A x - b^T x, \quad \Omega = \{x \in \mathbb{R}^n : Bx \leq c\}, \quad (11.3)$$

kde A je symetrická pozitivně definitní matice řádu n , matice $B = \{b_i\}$ je typu $r \times n$ a $c \in \mathbb{R}^r$. Předpokládáme, že $\Omega \neq \emptyset$.

V každém přípustném bodě x definujeme *aktivní množinu* následujícím předpisem

$$\mathcal{A}(x) = \{i \in \{1, \dots, r\} : b_i x = c_i\}.$$

Polyhedrální množiny

Abychom lépe porozuměli podmínkám řešitelnosti úlohám s omezeními ve tvaru nerovností (11.3), je užitečné zaměřit se na geometrii polyhedrálních množin. K tomu zavedeme některé nové pojmy.

Připomeňme, že množinu $\mathcal{C} \subseteq \mathbb{R}^n$ nazveme (konvexním) *kuželem*, jestliže $\alpha x \in \mathcal{C}$ pro všechny $\alpha \geq 0$ a $x \in \mathcal{C}$. Dále si zavedme pojem *polyhedrální kužel*, který definujeme předpisem

$$\mathcal{C} = \{x \in \mathbb{R}^n : Bx \leq o\},$$

kde $B \in \mathbb{R}^{m \times n}$ je daná matice. Všimněme si, že polyhedrální kužel je uzavřená konvexní množina.

Polyhedrální množinu Ω můžeme popsat jako součet konvexního obalu konečné množiny bodů a polyhedrálního kužele, jehož prvky jsou vektory z množiny Ω . Toto tvrzení zformulujeme v následující větě, kterou ponecháme bez důkazu.

Věta 11.1. *Množina $\Omega \subseteq \mathbb{R}^n$ je polyhedrální tehdy a jen tehdy, jestliže existuje neprázdná množina n -rozměrných vektorů $\{x_1, \dots, x_k\}$ a polyhedrální kužel $\mathcal{C} \subseteq \mathbb{R}^n$ takový, že platí*

$$\Omega = \mathcal{C} + \text{conv}\{x_1, \dots, x_k\}.$$

Farkasova věta

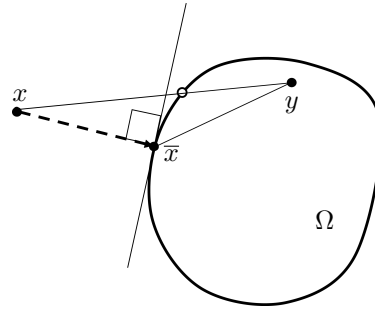
Abychom mohli převést geometrické podmínky minima pro úlohu s nerovnostními omezeními do algebraické podoby s Lagrangeovými multiplikatory, potřebujeme následující tvrzení - Farkasovu větu.

Předtím než tuto větu zformulujeme, budeme potřebovat ještě jedno tvrzení, které charakterizuje *projekci* P_Ω na (uzavřenou) konvexní množinu $\Omega \subset \mathbb{R}^n$ jako zobrazení, které přiřazuje každému $x \in \mathbb{R}^n$ nejbližší bod $\bar{x} \in \Omega$ (viz obrázek 11.2). O existenci tohoto bodu se vyjadřuje následující tvrzení (které necháme bez důkazu).

Věta 11.2. *Nechť $\Omega \subseteq \mathbb{R}^n$ je neprázdná uzavřená konvexní množina a $x \in \mathbb{R}^n$. Potom existuje právě jeden bod $\bar{x} \in \Omega$, který má minimální vzdálenost od x a pro každý $y \in \Omega$ platí*

$$(x - \bar{x})^T (y - \bar{x}) \leq 0. \tag{11.4}$$

Nyní už můžeme zformulovat Farkasovu větu



Obr. 11.2: Projekce na konvexní množinu

Věta 11.3. *Nechť $B \in \mathbb{R}^{m \times n}$ a $h \in \mathbb{R}^n$. Pak právě jedna z následujících úloh má řešení:*

- (I) Najdi $d \in \mathbb{R}^n$ takové, že $Bd \leq o$ a $h^T d > 0$.
 (II) Najdi $y \in \mathbb{R}^m$ takové, že $B^T y = h$ a $y \geq o$.

Důkaz. Nejprve předpokládáme, že úloha (II) má řešení. To znamená, že existuje $y \geq o$ takové, že $B^T y = h$. Nechť $d \in \mathbb{R}^n$ je takové, že $Bd \leq o$. Pak

$$h^T d = (B^T y)^T d = y^T B d \leq 0,$$

tedy úloha (I) nemá řešení.

Nyní předpokládejme, že úloha (II) nemá řešení a označme

$$\Omega = \{x \in \mathbb{R}^n : x = B^T y, y \geq o\}.$$

Náš předpoklad pak znamená, že $h \notin \Omega$. Označme symbolem $\hat{h} \in \Omega$ projekci h na množinu Ω a $d = h - \hat{h}$. Ze vztahu (11.4) plyne, že pro libovolné $x \in \Omega$ platí

$$d^T (x - \hat{h}) = (h - \hat{h})^T (x - \hat{h}) \leq 0$$

nebo

$$d^T x \leq d^T \hat{h} = \alpha.$$

Díky tomu, že $o \in \Omega$, platí $\alpha \geq 0$ a navíc dosazením $x = B^T y$, získáme pro libovolné $y \geq o$

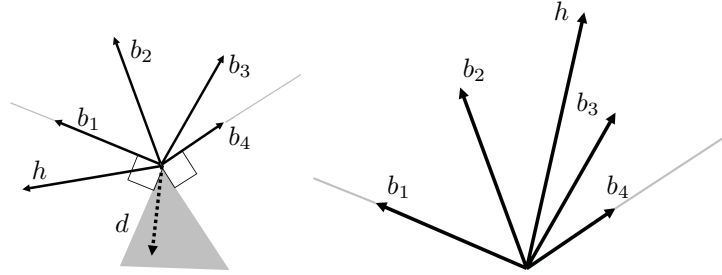
$$y^T B d = d^T B^T y \leq \alpha.$$

Protože složky y mohou být libovolně velké, platí $Bd \leq o$. To znamená, že d splňuje první podmínku úlohy (I). Abychom ověřili druhou podmínku, připomeňme, že podle předpokladu $h \notin \Omega$. Z toho plyne platnost $d \neq o$, dále $d^T (h - \hat{h}) = \|d\|^2 > 0$ a

$$d^T h > d^T \hat{h} = \alpha \geq 0.$$

Takže d je řešení úlohy (I). □

Tvrzení Farkasovy věty je nutno použít při důkazu KKT podmínek pro kvadratické úlohy s nerovnostními omezeními. Geometrická ilustrace Farkasovy věty je na obrázku 11.3.



Obr. 11.3: Farkasova věta: řešení úlohy (I) (vlevo) a (II) (vpravo)

11.3 Podmínky minima pro kvadratické programování s nerovnostmi

Úloha kvadratického programování s nerovnostmi (11.3) je výrazně komplikovanější než úloha kvadratického programování s rovnostmi (9.3). Hlavní otázky týkající se podmínek optimality a řešitelnosti (11.3) jsou zformulovány v následujícím tvrzení.

Věta 11.4. *Nechť je problém (11.3) definován pomocí symetrické matice $A \in \mathbb{R}^{n \times n}$, matice omezení $B \in \mathbb{R}^{r \times n}$ se sloupcovou hodnotí menší než n , a vektorů $b \in \mathbb{R}^n$, $c \in \mathcal{H}(B)$. Nechť \mathcal{C} označuje kužel přípustných směrů množiny Ω . Potom platí:*

- (i) *Jestliže úloha (11.3) má řešení, pak je $A|\mathcal{C}$ pozitivně semidefinitní.*
(ii) *Jestliže $\bar{x} \in \Omega$ je řešení (11.3), pak*

$$(A\bar{x} - b)^T d \geq 0 \quad (11.5)$$

pro každé přípustné d z množiny Ω v bodě \bar{x} .

- (iii) *Jestliže $\bar{x} \in \Omega$ je řešení (11.3), pak existuje vektor $\bar{\mu} \in \mathbb{R}^r$ tak, že*

$$\bar{\mu} \geq 0, \quad A\bar{x} - b + B^T \bar{\mu} = 0 \quad \text{a} \quad \bar{\mu}^T (B\bar{x} - c) = 0. \quad (11.6)$$

Důkaz. (i) Nechť \bar{x} je řešení problému (11.3), takže pro každé $d \in \mathbb{R}^n$, které splňuje $\bar{x} + \alpha d \in \Omega$, a $\alpha \in \mathbb{R}$

$$0 \leq f(\bar{x} + \alpha d) - f(\bar{x}) = \alpha(A\bar{x} - b)^T d + \frac{\alpha^2}{2} d^T A d. \quad (11.7)$$

Zvolíme-li $d \in \mathcal{C}$ a α dostatečně velké, dostaneme z nezápornosti pravé strany (11.7) že $d^T Ad \geq 0$. Tedy $A|_{\mathcal{C}}$ musí být pozitivně semidefinitní.

(ii) Necht \bar{x} je řešení problému (11.3) a d označuje přípustný směr z Ω v bodě \bar{x} , pak pravá strana

$$f(\bar{x} + \alpha d) - f(\bar{x}) = \alpha(A\bar{x} - b)^T d + \frac{\alpha^2}{2} d^T Ad \quad (11.8)$$

je nezáporná pro všechny dostatečně malé $\alpha > 0$. Abychom dokázali (11.5), stačí zvolit dostatečně malé $\alpha > 0$ tak, že z nezápornosti pravé strany (11.8) plyne

$$\alpha(A\bar{x} - b)^T d \geq 0.$$

(iii) Nejprve si všimněme, že jestliže \bar{x} je řešením (11.3), pak d je přípustný vektor množiny Ω v bodě \bar{x} tehdy a jen tehdy, pokud d je přípustný vektor množiny

$$\bar{\Omega} = \{x \in \mathbb{R}^n : B_{\mathcal{A}^*} x \leq c_{\mathcal{A}}\}$$

v bodě \bar{x} , kde $\mathcal{A} = \mathcal{A}(\bar{x})$ je aktivní množina v bodě \bar{x} . Takže podle (ii) je \bar{x} řešením také

$$\min_{x \in \bar{\Omega}} f(x), \quad \bar{\Omega} = \{x \in \mathbb{R}^n : B_{\mathcal{A}^*} x \leq c_{\mathcal{A}}\}.$$

Označme $h = -(A\bar{x} - b)$, z (ii) pak plyne, že úloha nalézt $d \in \mathbb{R}^n$ tak, že

$$B_{\mathcal{A}^*} d \leq o \quad \text{a} \quad h^T d > 0,$$

nemá řešení. Díky tomu, můžeme použít Farkasovu větu 11.3, abychom získali $y \in \mathbb{R}^m$ tak, že

$$(B_{\mathcal{A}^*})^T y = h_{\mathcal{A}} \quad \text{a} \quad y \geq o.$$

Označme $\bar{\mu} \in \mathbb{R}^m$ jako vektor získaný z vektoru y tak, že $[\bar{\mu}]_i = 0$ pro $i \notin \mathcal{A}$ a $\bar{\mu}_{\mathcal{A}} = y$. Je snadné ověřit, že vektor $\bar{\mu}$ splňuje podmínky (11.6). \square

Podmínky (11.6) nazýváme *KKT podmínkami pro úlohy s omezením ve tvaru nerovností*. Poslední z těchto podmínek, rovnost $\bar{\mu}^T (B\bar{x} - c) = 0$, nazýváme *podmínkou komplementarity*.

11.4 Nutné podmínky pro úlohy nelineárního programování

Vraťme se nyní k minimalizaci nelineární funkce s omezením ve tvaru nerovností (11.1). Abychom mohli zkoumat podmínku minima nelineární funkce s nelineárními omezeními, potřebujeme tečný kužel v bodě x k množině Ω , který jsme si zavedli v kapitole 9. Připomeňme, že jsme si tento tečný kužel v bodě \bar{x} k množině Ω zavedli jako množinu

$$T_{\Omega}(\bar{x}) = \{d \in \mathbb{R}^n : d = x'(0) \text{ pro přípustnou křivku } x \text{ množiny } \Omega\} \cup \{o\}.$$

Dále si potřebujeme zavést ještě dvě speciální množiny - kužel přípustných směrů a kužel směrů poklesu.

Definice 11.5. Kužel přípustných směrů v \bar{x} k množině $\Omega = \{x \in \mathbb{R}^n : g(x) \leq 0\}$ zavedeme jako množinu

$$V_{\Omega}(\bar{x}) = \{d \in \mathbb{R}^n : \nabla g_i(\bar{x})^T d \leq 0, \quad i \in \mathcal{A}(\bar{x})\}.$$

Kužel směrů poklesu funkce f v \bar{x} zavedeme jako množinu

$$D(\bar{x}) = \{d \in \mathbb{R}^n : \nabla f(\bar{x})^T d < 0\}.$$

Pomocí uvedených kuželů si zavedeme nutné podmínky minima pro úlohu (11.1). Nejprve využijeme pro podmínku minima kužel přípustných směrů a kužel směrů poklesu.

Věta 11.6. *Nechť problém (11.1) má lokální minimum v \bar{x} . Pak existuje vektor Lagrangeových multiplikátorů $\bar{\mu}$ se složkami $\bar{\mu}_j$, $j \in \mathcal{I} := \{1, \dots, r\}$ takovými, že následující systém*

$$\nabla_x L(\bar{x}, \bar{\mu}) = \nabla f(\bar{x}) + \nabla g(\bar{x})^T \bar{\mu} = 0,$$

$$\bar{\mu}_j \geq 0, \quad \forall j \in \mathcal{I},$$

$$\bar{\mu}_j \cdot g_j(\bar{x}) = 0, \quad \forall j \in \mathcal{I},$$

je splněn v $(\bar{x}, \bar{\mu})$ tehdy a jen tehdy, pokud platí podmínka

$$D(\bar{x}) \cap V_{\Omega}(\bar{x}) = \emptyset.$$

Nyní využijeme pro podmínku minima tečný kužel a kužel směrů poklesu.

Věta 11.7. *Nechť problém (11.1) má lokální minimum v \bar{x} . Pak platí podmínka*

$$D(\bar{x}) \cap T_{\Omega}(\bar{x}) = \emptyset.$$

Ze dvou předchozích vět můžeme v případě, že platí

$$T_{\Omega}(\bar{x}) = V_{\Omega}(\bar{x}),$$

zavést výsledné tvrzení.

Věta 11.8. *Nechť problém (11.1) má lokální minimum v \bar{x} . Jestliže v bodě \bar{x} platí*

$$T_{\Omega}(\bar{x}) = V_{\Omega}(\bar{x}),$$

pak existuje vektor Lagrangeových multiplikátorů $\bar{\mu}$ se složkami $\bar{\mu}_j$, $j \in \mathcal{I} := \{1 \text{ až } r\}$ takovými, že následující systém je splněn v $(\bar{x}, \bar{\mu})$:

$$\nabla_x L(\bar{x}, \bar{\mu}) = \nabla f(\bar{x}) + \nabla g(\bar{x})^T \bar{\mu} = o, \quad (11.9)$$

$$\bar{\mu}_j \geq 0, \quad \forall j \in \mathcal{I}, \quad (11.10)$$

$$\bar{\mu}_j \cdot g_j(\bar{x}) = 0, \quad \forall j \in \mathcal{I}. \quad (11.11)$$

Poznámka 11.9. Přípustný bod, který splňuje podmínku

$$T_{\Omega}(\bar{x}) = V_{\Omega}(\bar{x}),$$

bývá v některých textech označován jako *kvaziregulární bod*.

Podmínku optimality pro úlohu s nerovnostním omezením lze zavést i pomocí jiné geometrické konstrukce než té, která využívá tečný kužel. K tomu si musíme zavést další charakteristiku omezujících podmínek a tou je normálový kužel.

Definice 11.10. Množinu $N_{\Omega}(\bar{x})$ nazveme *normálový kužel* v \bar{x} k množině Ω , pokud je ortogonálním doplňkem tečného kuželu v \bar{x} k množině Ω , tzn.

$$N_{\Omega}(\bar{x}) = \{v : v^T w \leq 0 \quad \forall w \in T_{\Omega}(\bar{x})\}.$$

Pomocí normálového kužele můžeme podmínku optimality shrnout v následující větě.

Věta 11.11. *Nechť je problém (11.1) definován pomocí diferencovatelných zobrazení f a g . Nechť problém (11.1) má lokální minimum v \bar{x} , pak*

$$o \in \nabla f(\bar{x}) + N_{\Omega}(\bar{x}).$$

Podmínky Fritze - Johna

Nejprve si zavedeme obecné podmínky minima, které nepotřebují splnění žádných dodatečných podmínek na nerovnostní vazby.

Věta 11.12. *Nechť problém (11.1) má lokální minimum v \bar{x} . Pak existuje vektor Lagrangeových multiplikátorů $\bar{\mu}$ se složkami $\bar{\mu}_j$, $j \in \{0, \dots, r\}$ takovými, že následující systém je splněn v $(\bar{x}, \bar{\mu})$:*

$$\nabla f(\bar{x})\bar{\mu}_0 + \nabla g(\bar{x})^T \bar{\mu} = o, \quad (11.12)$$

$$\bar{\mu}_j \geq 0, \quad \forall j \in \{0, \dots, r\} \text{ a } \exists j \in \{0, \dots, r\} : \bar{\mu}_j \neq 0, \quad (11.13)$$

$$\bar{\mu}_j \cdot g_j(\bar{x}) = 0, \quad \forall j \in \{0, \dots, r\}. \quad (11.14)$$

Systém (11.12)-(11.14) s podmínkou $g(\bar{x}) \leq o$ nazýváme *Fritz-Johnovou podmínkou*.

KKT podmínky

Nyní si zavedeme podmínky minima, pro jejichž použití je nezbytné splnění další podmínky na nerovnostní vazby. Touto podmínkou je podmínka regularity.

Definice 11.13. Mějme minimalizační úlohu (11.1). Přípustný bod x nazveme *regulární*, jestliže gradienty aktivních nerovnostních vazeb jsou lineárně nezávislé.

Poznámka 11.14. Přípustný bod, který splňuje podmínku regularity, bývá v některých textech označován jako bod splňující klasifikaci omezení na lineární nezávislost¹.

Podmínky minima, které jsou speciálním případem Fritz-Johnovy podmínky, v případě, že bod \bar{x} je regulární, nazveme Karush-Kuhn-Tuckerovou podmínkou a zavedeme si ji v následujícím tvrzení. Toto tvrzení získáme snadno. Uvědomme si, že z regularity \bar{x} plyne, že pro

$$\nabla f(\bar{x})\bar{\mu}_0 + \nabla g(\bar{x})^T \bar{\mu} = o, \quad (11.15)$$

kde $\bar{\mu}_j$ ($j \in \{1, \dots, r\}$) nejsou všechny nulové, je $\bar{\mu}_0 \neq 0$. Nyní již stačí podělit poslední rovnost multiplikátorem $\bar{\mu}_0$ a dostaneme se k tvrzení následující věty.

Věta 11.15. *Nechť problém (11.1) má lokální minimum v \bar{x} a nechť je pro tento bod splněna podmínka regularity. Pak existuje jediný vektor Lagrangeových multiplikátorů $\bar{\mu}$ se složkami $\bar{\mu}_j$, $j \in \mathcal{I} := \{1, \dots, r\}$ takovými, že následující systém je splněn v $(\bar{x}, \bar{\mu})$:*

$$\nabla_x L(\bar{x}, \bar{\mu}) = \nabla f(\bar{x}) + \nabla g(\bar{x})^T \bar{\mu} = o, \quad (11.16)$$

$$\bar{\mu}_j \geq 0, \quad \forall j \in \mathcal{I}, \quad (11.17)$$

$$\bar{\mu}_j \cdot g_j(\bar{x}) = 0, \quad \forall j \in \mathcal{I}. \quad (11.18)$$

¹V anglické literatuře najde čtenář tuto podmínku pod názvem Linear Independence Constraint Qualification či pod zkratkou LICQ.

Dvojici $(\bar{x}, \bar{\mu})$ která splňuje (11.16)-(11.18) s podmínkou $g(\bar{x}) \leq 0$, budeme nazývat *Karush-Kuhn-Tuckerova dvojice* nebo stručně *KKT dvojice*. Systém (11.16)-(11.18) s podmínkou $g(\bar{x}) \leq 0$ nazýváme *Karush-Kuhn-Tuckerovou podmínkou* nebo stručně *KKT podmínkou*.

Poznámka 11.16. Je dobré si uvědomit, že podmínka regularity v přípustném bodě x nám zaručuje jednoznačnost λ splňující podmínky (11.16)-(11.18), ale v případě „větších“ úloh je značně komplikované ji ověřit, a proto je v tomto případě její použití omezené. V poslední části této kapitoly ukážeme, jak tuto podmínku nahradit „slabšími“ podmínkami tak, abychom měli zajištěnu alespoň existenci Lagrangeových multiplikátorů.

Nyní si uvedeme příklad, který nespĺňuje podmínku regularity v minimu.

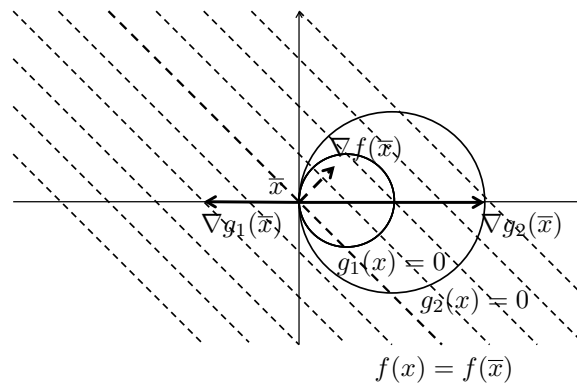
Příklad 11.17. Mějme následující minimalizační úlohu s nerovnostními omezeními

$$\min_{(x_1, x_2) \in \Omega} x_1 + x_2, \quad \Omega = \{(x_1, x_2) \in \mathbb{R}^2 : (x_1 - 1)^2 + x_2^2 \leq 1, \quad -(x_1 - 2)^2 - x_2^2 \leq -4\}.$$

Napište Fritz-Johnovu a Karush-Kuhn-Tuckerovu podmínku pro tuto úlohu a najděte dvojice, které je splňují.

Řešení. Pomůžeme si obrázkem 11.4, na kterém jsou znázorněny čárkovaně vrstevnice funkce f a nepřerušovanou čarou jsou znázorněny kružnice, jejichž průsečík představuje přípustnou množinu. Je snadné ukázat, že jediným přípustným bodem, který je řešením minimalizační úlohy, je bod o složkách

$$\bar{x}_1 = 0, \quad \bar{x}_2 = 0.$$



Obr. 11.4: Minimum lineární funkce na přípustné množině

Fritz-Johnovy podmínky pro tuto úlohu jsou ve tvaru

$$\mu_0 \nabla_{(x_1, x_2)} (x_1 + x_2) + \mu_1 \nabla_{(x_1, x_2)} ((x_1 - 1)^2 + x_2^2 - 1) +$$

$$\begin{aligned}
& +\mu_2 \nabla_{(x_1, x_2)} \left(-(x_1 - 2)^2 - x_2^2 + 4 \right) = 0, \\
& \mu_{0,1,2} \geq 0 \text{ a aspoň jedno z } \mu_{0,1,2} \text{ je nenulové,} \\
& \mu_1 \left((x_1 - 1)^2 + x_2^2 - 1 \right) = 0, \quad \mu_2 \left(-(x_1 - 2)^2 - x_2^2 + 4 \right) = 0, \\
& (x_1 - 1)^2 + x_2^2 \leq 1, \quad -(x_1 - 2)^2 - x_2^2 \leq -4.
\end{aligned}$$

Z tohoto systému dostaneme po derivování následující systém rovnic a nerovnic

$$\begin{aligned}
\mu_0 + 2\mu_1(x_1 - 1) - 2\mu_2(x_1 - 2) &= 0, \quad \mu_0 + 2\mu_1x_2 - 2\mu_2x_2 = 0, \\
\mu_{0,1,2} &\geq 0 \text{ a aspoň jedno z } \mu_{0,1,2} \text{ je nenulové,} \\
\mu_1 \left((x_1 - 1)^2 + x_2^2 - 1 \right) &= 0, \quad \mu_2 \left(-(x_1 - 2)^2 - x_2^2 + 4 \right) = 0, \\
(x_1 - 1)^2 + x_2^2 &\leq 1, \quad -(x_1 - 2)^2 - x_2^2 \leq -4.
\end{aligned}$$

Jejich řešením dostaneme multiplikátory

$$\mu_0 = 0, \quad \mu_1 = 2\mu_2 > 0.$$

Všimněme si, že $\mu_0 = 0$ a tedy (\bar{x}_1, \bar{x}_2) není regulární.

Karush-Kuhn-Tuckerovy podmínky pro tuto úlohu jsou ve tvaru

$$\begin{aligned}
& \nabla_{(x_1, x_2)} (x_1 + x_2) + \mu_1 \nabla_{(x_1, x_2)} \left((x_1 - 1)^2 + x_2^2 - 1 \right) + \\
& + \mu_2 \nabla_{(x_1, x_2)} \left(-(x_1 - 2)^2 - x_2^2 + 4 \right) = 0, \\
& \mu_{1,2} \geq 0, \\
& \mu_1 \left((x_1 - 1)^2 + x_2^2 - 1 \right) = 0, \quad \mu_2 \left(-(x_1 - 2)^2 - x_2^2 + 4 \right) = 0, \\
& (x_1 - 1)^2 + x_2^2 \leq 1, \quad -(x_1 - 2)^2 - x_2^2 \leq -4.
\end{aligned}$$

Z tohoto systému dostaneme po derivování následující systém rovnic a nerovnic

$$\begin{aligned}
1 + 2\mu_1(x_1 - 1) - 2\mu_2(x_1 - 2) &= 0, \quad 1 + 2\mu_1x_2 - 2\mu_2x_2 = 0, \\
\mu_{1,2} &\geq 0, \quad \mu_1 \left((x_1 - 1)^2 + x_2^2 - 1 \right) = 0, \quad \mu_2 \left(-(x_1 - 2)^2 - x_2^2 + 4 \right) = 0, \\
(x_1 - 1)^2 + x_2^2 &\leq 1, \quad -(x_1 - 2)^2 - x_2^2 \leq -4.
\end{aligned}$$

Pro tento systém neexistují multiplikátory μ_1, μ_2 , které by jej splnily.

Pro tuto úlohu je tedy splněna obecnější Fritz-Johnova podmínka, ale není splněna Karush-Kuhn-Tuckerova podmínka. \blacktriangle

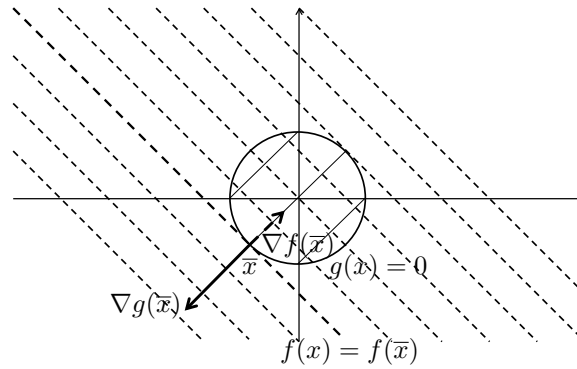
V následujících příkladech již bude podmínka regularity splněna a proto má smysl ověřovat splnění Karush-Kuhn-Tuckerovy podmínky.

Příklad 11.18. Mějme následující minimalizační úlohu s jedním nerovnostním omezením

$$\min_{(x_1, x_2) \in \Omega} x_1 + x_2, \quad \Omega = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 2\}.$$

Napište Karush-Kuhn-Tuckerovu podmínku pro tuto úlohu a najděte Karush-Kuhn-Tuckerovy dvojice, které ji splňují.

Řešení. Pomůžeme si obrázkem 11.5, na kterém jsou znázorněny čárkovaně vrstevnice funkce f a nepřerušovanou čarou je znázorněna kružnice, která představuje hranici přípustné množiny. Přípustná množina je vyšrafována plnými čarami. Minimum je dosaženo v bodě \bar{x} .



Obr. 11.5: Minimum lineární funkce na přípustné množině

Karush-Kuhn-Tuckerovy podmínky pro tuto úlohu jsou ve tvaru

$$\nabla_{(x_1, x_2)} (x_1 + x_2) + \mu \nabla_{(x_1, x_2)} (x_1^2 + x_2^2 - 2) = 0,$$

$$\mu \geq 0,$$

$$\mu(x_1^2 + x_2^2 - 2) = 0,$$

$$x_1^2 + x_2^2 \leq 2.$$

Z tohoto systému dostaneme po derivování následující systém rovnic a nerovnic

$$1 + 2\mu x_1 = 0, \quad 1 + 2\mu x_2 = 0,$$

$$\mu \geq 0, \quad \mu(x_1^2 + x_2^2 - 2) = 0,$$

$$x_1^2 + x_2^2 \leq 2.$$

Jejich řešením dostaneme složky řešení

$$\bar{x}_1 = \bar{x}_2 = -1, \quad \mu = 0,5.$$

Dá se ukázat, že řešením minimalizační úlohy je bod o složkách

$$\bar{x}_1 = \bar{x}_2 = -1.$$

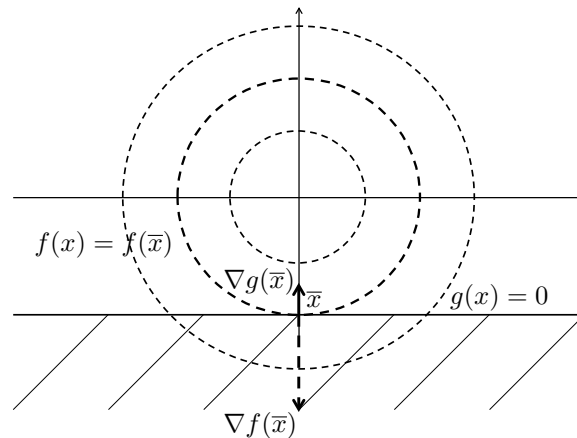
▲

Příklad 11.19. Mějme následující minimalizační úlohu s jedním nerovnostním omezením

$$\min_{(x_1, x_2) \in \Omega} x_1^2 + x_2^2, \quad \Omega = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \leq -5\}.$$

Napište Karush-Kuhn-Tuckerovu podmínku pro tuto úlohu a najděte Karush-Kuhn-Tuckerovy dvojice, které ji splňují.

Řešení. Pomůžeme si obrázkem 11.6, na kterém jsou znázorněny čárkované kruhové vrstevnice funkce f a nepřerušovanou čarou je znázorněna přímka, která představuje hranici přípustné množiny. Přípustná množina je vyšrafována plnými čarami. Minimum je dosaženo v bodě \bar{x} .



Obr. 11.6: Minimum kvadratické funkce na přípustné množině

Karush-Kuhn-Tuckerovy podmínky pro tuto úlohu jsou ve tvaru

$$\begin{aligned} \nabla_{(x_1, x_2)} (x_1^2 + x_2^2) + \mu \nabla_{(x_1, x_2)} (x_2 + 5) &= 0, \\ \mu &\geq 0, \\ \mu(x_2 + 5) &= 0, \\ x_2 &\leq -5. \end{aligned}$$

Z tohoto systému dostaneme po derivování následující systém rovnic a nerovnic

$$\begin{aligned} 2x_1 &= 0, & 2x_2 + \mu &= 0, \\ \mu &\geq 0, & \mu(x_2 + 5) &= 0, \end{aligned}$$

$$x_2 \leq -5.$$

Jejich řešením dostaneme složky řešení

$$\bar{x}_1 = 0, \bar{x}_2 = -5, \mu = 10.$$

Dá se ukázat, že řešením minimalizační úlohy je bod o složkách

$$\bar{x}_1 = 0, \bar{x}_2 = -5.$$

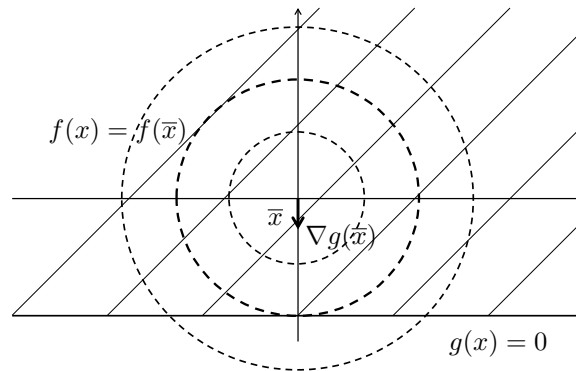
▲

Příklad 11.20. Mějme následující minimalizační úlohu s jedním nerovnostním omezením

$$\min_{(x_1, x_2) \in \Omega} x_1^2 + x_2^2, \quad \Omega = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \geq -5\}.$$

Napište Karush-Kuhn-Tuckerovu podmínku pro tuto úlohu a najděte Karush-Kuhn-Tuckerovy dvojice, které ji splňují.

Řešení. Pomůžeme si obrázkem 11.7, na kterém jsou znázorněny čárkované kruhové vrstevnice funkce f a nepřerušovanou čarou je znázorněna přímka, která představuje hranici přípustné množiny. Přípustná množina je vyšrafována plnými čarami. Minimum je dosaženo v bodě \bar{x} .



Obr. 11.7: Minimum kvadratické funkce na přípustné množině

Karush-Kuhn-Tuckerovy podmínky pro tuto úlohu jsou ve tvaru

$$\nabla_{(x_1, x_2)} (x_1^2 + x_2^2) + \mu \nabla_{(x_1, x_2)} (-x_2 - 5) = 0,$$

$$\mu \geq 0,$$

$$\mu(-x_2 - 5) = 0,$$

$$x_2 \geq -5.$$

Z tohoto systému dostaneme po derivování následující systém rovnic a nerovnic

$$\begin{aligned} 2x_1 &= 0, & 2x_2 - \mu &= 0, \\ \mu &\geq 0, & \mu(-x_2 - 5) &= 0, \\ x_2 &\geq -5. \end{aligned}$$

Jejich řešením dostaneme složky řešení

$$\bar{x}_1 = 0, \quad \bar{x}_2 = 0, \quad \mu = 0.$$

Dá se ukázat, že řešením minimalizační úlohy je bod o složkách

$$\bar{x}_1 = 0, \quad \bar{x}_2 = 0.$$

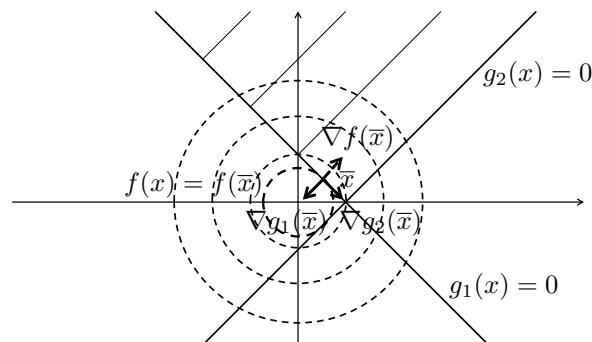
▲

Příklad 11.21. Mějme následující minimalizační úlohu se dvěma nerovnostními omezeními

$$\min_{(x_1, x_2) \in \Omega} x_1^2 + x_2^2, \quad \Omega = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 \geq 1, \quad x_1 - x_2 \leq 1\}.$$

Napište Karush-Kuhn-Tuckerovu podmínku pro tuto úlohu a najděte Karush-Kuhn-Tuckerovy dvojice, které ji splňují.

Řešení. Pomůžeme si obrázkem 11.8, na kterém jsou znázorněny čárkovaně kruhové vrstevnice funkce f a nepřerušovanou čarou jsou znázorněny přímky, jejichž průsečík představuje přípustnou množinu. Přípustná množina je vyšrafována plnými čarami. Minimum je dosaženo v bodě \bar{x} .



Obr. 11.8: Minimum kvadratické funkce na přípustné množině

Karush-Kuhn-Tuckerovy podmínky pro tuto úlohu jsou ve tvaru

$$\begin{aligned} \nabla_{(x_1, x_2)} (x_1^2 + x_2^2) + \mu_1 \nabla_{(x_1, x_2)} (-x_1 - x_2 + 1) + \mu_2 \nabla_{(x_1, x_2)} (x_1 - x_2 - 1) &= 0, \\ \mu_{1,2} &\geq 0, \\ \mu_1(-x_1 - x_2 + 1) &= 0, \quad \mu_2(x_1 - x_2 - 1) = 0, \\ x_1 + x_2 &\geq 1, \quad x_1 - x_2 \leq 1. \end{aligned}$$

Z tohoto systému dostaneme po derivování následující systém rovnic a nerovnic

$$\begin{aligned} 2x_1 - \mu_1 + \mu_2 &= 0, \quad 2x_2 - \mu_1 - \mu_2 = 0, \\ \mu_{1,2} &\geq 0, \quad \mu_1(-x_1 - x_2 + 1) = 0, \quad \mu_2(x_1 - x_2 - 1) = 0, \\ x_1 + x_2 &\geq 1, \quad x_1 - x_2 \leq 1. \end{aligned}$$

Jejich řešením dostaneme složky řešení

$$\bar{x}_1 = \bar{x}_2 = 0,5, \quad \mu_1 = 1, \quad \mu_2 = 0.$$

Dá se ukázat, že řešením minimalizační úlohy je bod o složkách

$$\bar{x}_1 = \bar{x}_2 = 0,5.$$

▲

11.5 Klasifikace omezení pro úlohy s omezeními ve tvaru rovností a nerovností

Nyní se budeme zabývat úlohou optimalizace s omezeními ve tvaru rovností i nerovností

$$\min_{x \in \Omega} f(x), \quad \Omega = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0\}, \quad (11.19)$$

kde $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^r$, $m < n$, $r < n$. Budeme při tom opět předpokládat, že f je dvakrát spojitě diferencovatelná funkce a h a g jsou spojitě diferencovatelné.

Pro tento případ můžeme zobecnit definici kužele přípustných směrů.

Definice 11.22. Kužel přípustných směrů v \bar{x} k množině $\Omega = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0\}$ zavedeme jako množinu

$$V_{\Omega}(\bar{x}) = \{d \in \mathbb{R}^n : \nabla h_i(\bar{x})^T d = 0, \nabla g_j(\bar{x})^T d \leq 0, i \in 1, \dots, m, j \in \mathcal{A}(\bar{x})\}.$$

Pro úlohu s omezeními ve tvaru rovností a nerovností s využitím právě uvedené definice kužele přípustných směrů můžeme zobecnit tvrzení věty 11.8 a tím získáme tvrzení následující věty.

Věta 11.23. *Nechť problém (11.19) má lokální minimum v \bar{x} . Jestliže v bodě \bar{x} platí*

$$T_{\Omega}(\bar{x}) = V_{\Omega}(\bar{x}),$$

pak existuje vektor Lagrangeových multiplikátorů $\bar{\mu}$ se složkami $\bar{\mu}_j$, $j \in \mathcal{I} = \{0 \text{ až } r\}$ a $\bar{\lambda}$ se složkami $\bar{\lambda}_i$, $i \in \{1, \dots, m\}$ takovými, že následující systém je splněn v $(\bar{x}, \bar{\lambda}, \bar{\mu})$:

$$\nabla_x L(\bar{x}, \bar{\lambda}, \bar{\mu}) = \nabla f(\bar{x}) + \nabla h(\bar{x})^T \bar{\lambda} + \nabla g(\bar{x})^T \bar{\mu} = o,$$

$$\bar{\mu}_j \geq 0, \quad \forall j \in \mathcal{I},$$

$$\bar{\mu}_j \cdot g_j(\bar{x}) = 0, \quad \forall j \in \mathcal{I}.$$

Poznámka 11.24. Přípustný bod, který splňuje podmínku

$$T_{\Omega}(\bar{x}) = V_{\Omega}(\bar{x}),$$

bývá v některých textech označován jako kvaziregulární bod i pro množinu Ω z úlohy (11.19).

Pro naši obecnější úlohu je možné také zobecnit podmínku optimality obsaženou ve větě 11.11, která je formulována pomocí normálového kužele k množině Ω . Definice tohoto kužele je uvedena v definici 11.10. Toto zobecnění je obsaženo v následující větě.

Věta 11.25. *Nechť je problém (11.19) definován pomocí diferencovatelných zobrazení f, h a g . Nechť problém (11.19) má lokální minimum v \bar{x} , pak*

$$o \in \nabla f(\bar{x}) + N_{\Omega}(\bar{x}).$$

Pro tento obecnější případ také můžeme přeformulovat věty 11.12 a 11.15. K tomu potřebujeme nejdříve zavést opět podmínku regularity (nyní pro rovnostní a nerovnostní vazby).

Definice 11.26. Mějme minimalizační úlohu (11.19). Přípustný bod x nazveme *regulární*, jestliže gradienty rovnostních vazeb a gradienty aktivních nerovnostních vazeb jsou lineárně nezávislé.

Nyní můžeme přeformulovat věty 11.12 a 11.15 následujícím způsobem.

Věta 11.27. *Nechť problém (11.19) má lokální minimum v \bar{x} . Pak existuje vektor Lagrangeových multiplikátorů $\bar{\mu}$ se složkami $\bar{\mu}_j$, $j \in \{0, \dots, r\}$ a $\bar{\lambda}$ se složkami $\bar{\lambda}_i$, $i \in \{1, \dots, m\}$ takovými, že následující systém je splněn v $(\bar{x}, \bar{\lambda}, \bar{\mu})$:*

$$\nabla f(\bar{x})\bar{\mu}_0 + \nabla h(\bar{x})^T \bar{\lambda} + \nabla g(\bar{x})^T \bar{\mu} = o, \quad (11.20)$$

$$\bar{\mu}_j \geq 0, \quad \forall j \in \{0, \dots, r\} \text{ a } \bar{\mu}_0, \bar{\lambda}_1, \dots, \bar{\lambda}_m, \bar{\mu}_1, \dots, \bar{\mu}_r \text{ nejsou všechny rovny } 0, \quad (11.21)$$

$$\bar{\mu}_j \cdot g_j(\bar{x}) = 0, \quad \forall j \in \mathcal{I} = \{0, \dots, r\}. \quad (11.22)$$

Systém (11.20)-(11.22) s podmínkami $g(\bar{x}) \leq o$ a $h(\bar{x}) = o$ podobně jako u úlohy pouze s nerovnostními omezeními nazýváme *Fritz-Johnovou podmínkou*.

Věta 11.28. *Nechť problém (11.19) má lokální minimum v \bar{x} a nechť je pro tento bod splněna podmínka regularity. Pak existuje jediný vektor Lagrangeových multiplikátorů $\bar{\mu}$ se složkami $\bar{\mu}_j$, $j \in \mathcal{I} := \{1, \dots, r\}$ a $\bar{\lambda}$ se složkami $\bar{\lambda}_i$, $i \in \{1, \dots, m\}$ takovými, že následující systém je splněn v $(\bar{x}, \bar{\lambda}, \bar{\mu})$:*

$$\nabla_x L(\bar{x}, \bar{\lambda}, \bar{\mu}) = \nabla f(\bar{x}) + \nabla h(\bar{x})^T \bar{\lambda} + \nabla g(\bar{x})^T \bar{\mu} = o, \quad (11.23)$$

$$\bar{\mu}_j \geq 0, \quad \forall j \in \mathcal{I}, \quad (11.24)$$

$$\bar{\mu}_j \cdot g_j(\bar{x}) = 0, \quad \forall j \in \mathcal{I}. \quad (11.25)$$

Dvojici $(\bar{x}, (\bar{\lambda}, \bar{\mu}))$ která splňuje (11.23)-(11.25) s podmínkami $g(\bar{x}) \leq o$ a $h(\bar{x}) = o$, budeme podobně jako u úlohy pouze s nerovnostními omezeními nazývat *Karush-Kuhn-Tuckerova dvojice* nebo stručně *KKT dvojice*. Systém (11.23)-(11.25) s podmínkami $g(\bar{x}) \leq o$ a $h(\bar{x}) = o$ nazýváme *Karush-Kuhn-Tuckerovou podmínkou* nebo stručně *KKT podmínkou*.

Nyní se seznámíme s jinými podmínkami, které ve větě 11.28 nahradí podmínku regularity tak, aby byla zajištěna existence Lagrangeových multiplikátorů, které splňují KKT podmínku. Těmito dalším podmínkám i podmínce regularity budeme říkat klasifikace omezení¹. Tyto jiné podmínky jsou „slabší“ než podmínka regularity a zaručují pouze existenci a nikoliv jednoznačnost Lagrangeových multiplikátorů splňujících KKT podmínku.

¹V anglické literatuře najde čtenář tyto podmínky pod názvem Constraint Qualifications či pod zkratkou CQ.

První z těchto klasifikací omezení je klasifikace omezení na linearitu/konkávnost.

Věta 11.29. (O klasifikaci omezení na linearitu/konkávnost)^a Necht problém (11.19) má lokální minimum v \bar{x} a necht je splněna následující podmínka.

Složky h_i jsou lineární a složky g_j jsou konkávní.

Pak existuje vektor Lagrangeových multiplikátorů $\bar{\lambda}$ se složkami $\bar{\lambda}_i$, $i \in \mathcal{E} := \{1 \text{ až } m\}$ a $\bar{\mu}$ se složkami $\bar{\mu}_i$, $i \in \mathcal{I} := \{1, \dots, r\}$ takovými, že $(\bar{x}, \bar{\lambda}, \bar{\mu})$ je KKT dvojicí pro problém (11.19).

^aV anglické literatuře najde čtenář tuto podmínku pod názvem Linear/Concave Constraint Qualification.

Další z těchto klasifikací omezení je Mangasarian-Fromovitzova klasifikace omezení.

Věta 11.30. (O Mangasarian-Fromovitzově klasifikaci omezení)^a Necht problém (11.19) má lokální minimum v \bar{x} a necht jsou pro tento bod splněny následující podmínky.

Gradientsy $\nabla h_i(\bar{x})$, $i \in \mathcal{E} := \{1, \dots, m\}$ jsou lineárně nezávislé a existuje vektor d takový, že

$$\nabla h_i(\bar{x})^T d = 0, \quad \forall i \in \mathcal{E}, \quad \nabla g_j(\bar{x})^T d < 0, \quad \forall j \in \mathcal{A}(\bar{x}).$$

Pak existuje vektor Lagrangeových multiplikátorů $\bar{\lambda}$ se složkami $\bar{\lambda}_i$, $i \in \mathcal{E}$ a $\bar{\mu}$ se složkami $\bar{\mu}_i$, $i \in \mathcal{I}$ takovými, že $(\bar{x}, \bar{\lambda}, \bar{\mu})$ je KKT dvojicí pro problém (11.19).

^aV anglické literatuře najde čtenář tuto podmínku pod názvem Mangasarian-Fromovitz Constraint Qualification či pod zkratkou MFCQ.

Na závěr si uvedeme Slaterovu klasifikaci omezení.

Věta 11.31. (*O Slaterově klasifikaci omezení*)^a Necht problém (11.19) má lokální minimum v \bar{x} a necht jsou pro tento bod splněny následující podmínky.

Složky h_i jsou lineární a g_j jsou konvexní. Dále existuje přípustný bod x takový, že

$$g_j(x) < 0, \quad \forall j \in \mathcal{A}(\bar{x}).$$

Pak existuje vektor Lagrangeových multiplikátorů $\bar{\lambda}$ se složkami $\bar{\lambda}_i$, $i \in \mathcal{E}$ a $\bar{\mu}$ se složkami $\bar{\mu}_i$, $i \in \mathcal{I}$ takovými, že $(\bar{x}, \bar{\lambda}, \bar{\mu})$ je KKT dvojicí pro problém (11.19).

^aV anglické literatuře najde čtenář tuto podmínku pod názvem Slater Constraint Qualification či pod zkratkou SCQ.

Kapitola 12

Metody řešení úloh s nerovnostmi

V této kapitole se seznámíme se základními algoritmy pro nalezení minima nelineární funkce vzhledem k omezením ve tvaru nerovností. Seznámíme se postupně se třemi metodami řešení této úlohy. V klasické analýze se řešení úloh minimalizace s nerovnostními vazbami nazývají také *podmíněné extrémy*. Pro jednoduchost budeme předpokládat, že f je konvexní a že úloha má jediné řešení.

12.1 Metoda vnitřního bodu - bariéry

První jednoduchou možností, jak řešit úlohu s nerovnostními vazbami, je metoda bariéry. Ta využívá tzv. bariérovou funkci, kterou si zavedeme dále v textu.

Definice 12.1. Definujme pro množinu Ω z úlohy (11.1) funkci $b : \mathbb{R}^n \rightarrow \mathbb{R}$, která splňuje $b|_{\Omega} \geq 0$ a $b(x) \rightarrow +\infty$ pro $g(x) \rightarrow 0$. Takovou funkci b nazveme *bariérovou funkcí* příslušnou přípustné množině Ω .

Bariérovou funkci můžeme obecně zavést takto $b := \sum_{i=1}^m \varphi(g_i(x))$, kde $\varphi(y) \geq 0$ pro $y \leq 0$ a $\varphi(y) \rightarrow +\infty$ pro $y \rightarrow 0_-$.

Příklad 12.2. Sestrojte bariérovou funkci příslušnou přípustné množině

$$\Omega = \{x \in \mathbb{R}^n : g(x) \leq o\}, \quad g : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad m < n.$$

Řešení. Příkladem bariérové funkce příslušnou množině Ω jsou funkce

$$b(x) := - \sum_{i=1}^m \frac{c_i}{g_i(x)}$$

či

$$b(x) := - \sum_{i=1}^m c_i \ln(-g_i(x)),$$

kde $c_i \in \mathbb{R}$ jsou vhodné konstanty. ▲

Pomocí bariérové funkce můžeme úlohu s omezením ve tvaru nerovností (11.1) převést na úlohu bez omezení

$$\min_{\mu \geq 0} \inf_{x \in \mathbb{R}^n} f(x) + \mu b(x).$$

Není těžké si uvědomit, že řešení této pomocné úlohy leží vždy uvnitř přípustné množiny. Proto se tato *bariérová metoda* často označuje jako tzv. *metoda vnitřního bodu*.

Příklad 12.3. Mějme následující minimalizační úlohu s jedním nerovnostním omezením

$$\min_{x \in \Omega} x, \quad \Omega = \{x \in \mathbb{R} : -x + 1 \leq 0\}.$$

Sestrojte bariérovou funkci pro tuto úlohu a převedte tuto úlohu na úlohu bez omezení.

Řešení. Příkladem bariérové funkce příslušnou množině $\Omega = \{x \in \mathbb{R} : -x + 1 \leq 0\}$ je funkce

$$b(x) := -\frac{1}{-x + 1}.$$

S využitím této bariérové funkce můžeme přejít od úlohy s omezením $-x + 1 \leq 0$ k úloze bez omezení

$$\min_{\mu \geq 0} \inf_{x \in \mathbb{R}} \left\{ x - \frac{\mu}{-x + 1} : -x + 1 \leq 0 \right\}.$$

▲

Uvedme si bez důkazu následující tvrzení.

Věta 12.4. *Nechť $\mu \rightarrow 0_+$, pak řešení úlohy*

$$\inf_{x \in \mathbb{R}^n} \{f(x) + \mu b(x) : g(x) \leq 0\}$$

konverguje k řešení úlohy (11.1).

Toto tvrzení nás může inspirovat k následujícímu algoritmu metody bariéry pro řešení úlohy (11.1).

Algoritmus 12.1 **Metoda bariéry.**

Krok 0. {Inicializace.}
 Zvolme $\varepsilon > 0$ (přesnost), $x^0 \in \Omega = \{x \in \mathbb{R}^n : g(x) \leq 0\}$, $\mu > 0$,
 $\beta \in (0, 1)$, $k = 1$

Krok 1. {Základní cyklus pro metodu bariéry.}
for $k = 0, 1, 2, \dots$
 $x^k = \arg \min_{x \in \mathbb{R}^n} f(x) + \mu b(x)$
 if $\mu b(x^k) < \varepsilon$
 break
 else
 $\mu = \beta \mu$
 end
 $k = k + 1$
end for

Krok 2. {Řešení.}
 $\bar{x} = x^k$

12.2 Metoda aktivních (pracovních) množin - vnější penalta

Další velmi jednoduchá metoda převedení úlohy s omezením (11.1) spočívá v zahrnutí nerovnostních vazeb do cenové funkce tak, že k ní přičteme vhodný člen, který penalizuje porušení vazeb. Omezíme se zde, podobně jako v kapitole 10, na nejčastěji používanou metodu *kvadratické penalty*, která aproximuje řešení \bar{x} úlohy (11.1) řešením \bar{x}_ϱ úlohy

$$\min_{x \in \mathbb{R}^n} f_\varrho(x), \quad f_\varrho(x) = f(x) + \frac{1}{2} \varrho \|\alpha(x)\|^2, \quad \alpha(x) = \max\{g(x), 0\},$$

kde $\varrho > 0$ je penalizační parametr a $\|\alpha(x)\|^2$ je penalizační funkce. Funkci \max aplikujeme po složkách.

Intuitivně je zřejmé, že když je penalizační parametr ϱ velký, tak řešení \bar{x}_ϱ , ve kterém je dosaženo minimum penalizované funkce f_ϱ , nemůže být daleko od přípustné množiny. Je dokonce zřejmé, že kdyby $\varrho = \infty$, pak by bylo minimum f_ϱ řešením původní úlohy (11.1). Můžeme tedy očekávat, že pro dostatečně velké hodnoty penalizačního parametru ϱ bude řešení \bar{x} úlohy (11.1) blízko \bar{x}_ϱ . Je zřejmé, že řešení penalizované úlohy je typicky blízko přípustné množiny, avšak nepatří do ní. Proto se naše penalizační metoda nazývá také *metoda externí penalty*.

12.3 SMALBE

Následující algoritmus je modifikací algoritmu SMALE ze sekce 10.5. Jediným rozdílem je, že algoritmus SMALBE řeší ve vnitřní smyčce úlohu s jednoduchými omezeními zdola s přesností, která je řízena eukleidovskou normou projektovaného gradientu¹. Algoritmus SMALBE je (podobně jako SMALE) určen pro řešení problémů kvadratického programování, konkrétně pro úlohu

$$\min_{x \in \Omega} f(x), \quad f(x) = \frac{1}{2}x^T A x - b^T x, \quad \Omega = \{x \in \mathbb{R}^n : Bx = o, x \geq \ell\}.$$

Celý algoritmus pak vypadá následovně.

Algoritmus 12.2 Metoda semimonotónních rozšířených Lagrangiánů - Semimonotonic augmented Lagrangians (SMALBE).

Dána SPD matice $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times n}$, $\ell \in \mathbb{R}^n$.

Krok 0. {Inicializace.}

Zvolte $\eta > 0$, $\beta > 1$, $M > 0$, $\varrho_0 > 0$, $\lambda^0 \in \mathbb{R}^m$

for $k = 0, 1, 2, \dots$

Krok 1. {Vnitřní iterace s adaptivní kontrolou přesnosti.}

Nalezněte $x^k \geq \ell$ tak, aby platilo

$$\|g^P(x^k, \lambda^k, \varrho_k)\| \leq \min\{M\|Bx^k\|, \eta\}$$

Krok 2. {Aktualizace Lagrangeových multiplikátorů.}

$$\lambda^{k+1} = \lambda^k + \varrho_k Bx^k$$

Krok 3. {Aktualizace ϱ v případě, že růst Lagrangiánu není dostatečný.}

if $k > 0$ a $L(x^k, \lambda^k, \varrho_k) < L(x^{k-1}, \lambda^{k-1}, \varrho_{k-1}) + \frac{\varrho_k}{2}\|Bx^k\|^2$

$\varrho_{k+1} = \beta \varrho_k$

else

$\varrho_{k+1} = \varrho_k$

end if

end for

Příklady k procvičení

1. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = x_1^4 + x_1^2 + x_2^2.$$

¹Projektovaným gradientem myslíme gradient „projektovaný na“ přípustnou množinu $\Omega = \{x \in \mathbb{R}^n : Bx = c \wedge x \geq \ell\}$. Tento gradient označíme jako g^P .

Řešte úlohu optimalizace s omezením na nerovnost

$$\min_{x \in \Omega} f(x_1, x_2), \quad \Omega = \{x \in \mathbb{R}^2 : -x_1 - 3x_2 + 5 \leq 0\}.$$

Uvedenou úlohu s omezením na nerovnost řešte pomocí metody bariéry a vnější penalty.

2. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = x_1^4 + x_1^2 + x_2^2.$$

Řešte úlohu optimalizace s omezeními na nerovnost

$$\min_{x \in \Omega} f(x_1, x_2), \quad \Omega = \{x \in \mathbb{R}^2 : -2x_1 - 3x_2 - 1 \leq 0, -x_1 + x_2 + 5 \leq 0\}.$$

Uvedenou úlohu s omezeními na nerovnost řešte pomocí metody bariéry a vnější penalty.

3. Minimum kvadratické funkce

$$f(x) = \frac{1}{2}x^T Ax - b^T x$$

aproximuje (až na jednotky) průhyb struny zatížené jednotkovou silou a uchycené na obou koncích ($x_1 = x_n = 0$), kde $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$ a

$$A = (n-1) \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$b = -\frac{1}{n-1} \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 1 \\ 0 \end{bmatrix}.$$

Minimum kvadratické funkce f za podmínky $x_i \geq \ell_i$, kde $\ell_i < 0$ definuje překážku, popisuje průhyb struny nad tuhou překážkou. Otestujte metodu vnější penalty na problémech s překážkou určenou hodnotou $\tau < 0$.

Kapitola 13

Dualita v konvexním programování

V této kapitole si ukážeme alternativní formulace konvexních úloh kvadratického programování, jejichž vzájemný vztah je charakterizován pojmem dualita. Připomeňme si, že podle Wikipedie „a duality (in mathematics) translates concepts, theorems or mathematical structures into other concepts, theorems or structures, in a one-to-one fashion, often involution“. Dualita nás zajímá zejména proto, že často umožňuje přeformulovat původní úlohu na úlohu, jejíž řešení je mnohem snazší.

13.1 Dualita pro kvadratické programování s rovnostními vazbami

Základní myšlenku si nejprve vyložíme na úloze konvexního kvadratického programování s rovnostmi

$$\min_{x \in \Omega} f(x), \quad f(x) = \frac{1}{2}x^T Ax - b^T x, \quad \Omega = \{x \in \mathbb{R}^n : Bx = c\}, \quad (13.1)$$

kde A je symetrická pozitivně definitní a B má plnou řádkovou hodnost. Z našeho předpokladu plyne, že příslušná KKT soustava

$$\nabla_x L_0(x, \lambda) = Ax - b + B^T \lambda = o, \quad (13.2)$$

$$\nabla_\lambda L_0(x, \lambda) = Bx - c = o \quad (13.3)$$

má jediné řešení $(\bar{x}, \bar{\lambda})$, které můžeme najít tak, že si nejprve vyjádříme x z (13.2) a následně dosadíme za x do (13.3), takže dostaneme rovnici pro $\bar{\lambda}$. Tyto dva kroky si nyní spojíme se dvěma optimalizačními problémy.

Nejprve si všimněme, že rovnici (13.2) můžeme považovat za podmínku minima, jejíž splnění zaručuje, že

$$L_0(x, \lambda) = \frac{1}{2}x^T Ax - b^T x + \lambda^T (Bx - c)$$

dosahuje v x minimum vzhledem k x . Pro dané $\lambda \in \mathbb{R}^m$ je tedy *první krok ekvivalentní nalezení argumentu minima*

$$x = x(\lambda) = A^{-1}(b - B^T \lambda)$$

Lagrangiánu $L_0(x, \lambda)$ vzhledem k x . Pomocí tohoto pozorování si můžeme vyjádřit explicitně *duální funkci*

$$\begin{aligned} \Theta(\lambda) &= \inf_{x \in \mathbb{R}^n} L_0(x, \lambda) = \min_{x \in \mathbb{R}^n} L_0(x, \lambda) = L_0(x(\lambda), \lambda) \\ &= -\frac{1}{2}\lambda^T BA^{-1}B^T \lambda + (BA^{-1}b - c)^T \lambda - \frac{1}{2}b^T A^{-1}b \end{aligned}$$

a její gradient

$$\nabla \Theta(\lambda) = -BA^{-1}B^T \lambda + (BA^{-1}b - c). \quad (13.4)$$

Abychom našli interpretaci druhého kroku, dosaďme si za $x = x(\lambda)$ do (13.3), čímž dostaneme

$$-BA^{-1}B^T \lambda + (BA^{-1}b - c) = o.$$

Porovnáme-li levou stranu poslední rovnice s explicitním výrazem (13.4) pro $\nabla \Theta(\lambda)$, můžeme si poslední rovnici zapsat ve tvaru

$$\nabla \Theta(\lambda) = o.$$

Jelikož matice $BA^{-1}B^T$, což je Hessián funkce $-\Theta$, je podle předpokladu pozitivně definitní, dostáváme tak, že poslední výraz je ekvivalentní podmínce minima (2.7) pro argument minima $-\Theta$, což je současně argument maxima Θ . Proto KKT dvojice $(\bar{x}, \bar{\lambda})$ pro problém (13.1) s rovnostními vazbami řeší též *min-max problém*

$$L_0(\bar{x}, \bar{\lambda}) = \max_{\lambda \in \mathbb{R}^m} \min_{x \in \mathbb{R}^n} L_0(x, \lambda), \quad (13.5)$$

$\bar{\lambda}$ řeší *duální problém*

$$\Theta(\bar{\lambda}) = \max_{\lambda \in \mathbb{R}^m} \Theta(\lambda),$$

a protože \bar{x} je přípustný vektor, platí

$$f(\bar{x}) = L_0(\bar{x}, \lambda) = L_0(\bar{x}, \bar{\lambda}) = \Theta(\bar{\lambda}). \quad (13.6)$$

Navíc zřejmě platí

$$f(\bar{x}) = L_0(\bar{x}, \bar{\lambda}) = \min_{x \in \mathbb{R}^n} L_0(x, \bar{\lambda}) \leq L_0(x, \bar{\lambda}), \quad x \in \mathbb{R}^n. \quad (13.7)$$

Existuje ještě jeden ekvivalentní problém související s penalizační metodou. Jelikož

$$\sup_{\lambda \in \mathbb{R}^m} L_0(x, \lambda) = \infty \text{ pro } x \notin \Omega \quad \text{a} \quad \sup_{\lambda \in \mathbb{R}^m} L_0(x, \lambda) = f(x) \text{ pro } x \in \Omega,$$

platí, že řešení \bar{x} KKT soustavy (13.2)-(13.3) splňuje

$$f(\bar{x}) = \min_{x \in \mathbb{R}^n} \sup_{\lambda \in \mathbb{R}^m} L_0(x, \lambda). \quad (13.8)$$

Porovnáním (13.8) s (13.5) a (13.6) dostaneme *relaci duality*

$$\max_{\lambda \in \mathbb{R}^m} \min_{x \in \mathbb{R}^n} L_0(x, \lambda) = \min_{x \in \mathbb{R}^n} \sup_{\lambda \in \mathbb{R}^m} L_0(x, \lambda).$$

Použijeme-li (13.6) a (13.7), dostaneme, že $(\bar{x}, \bar{\lambda})$ řeší sedlobodový problém najít $(\bar{x}, \bar{\lambda})$ tak, aby pro každé $x \in \mathbb{R}^n$ a $\lambda \in \mathbb{R}^m$ platilo

$$L_0(\bar{x}, \lambda) \leq L_0(\bar{x}, \bar{\lambda}) \leq L_0(x, \bar{\lambda}). \quad (13.9)$$

Odvodili jsme si tak dva problémy *bez omezení*, které jsou ekvivalentní původnímu problému s rovnostními omezeními (13.1). Sedlobodová formulace obsahuje explicitně Lagrangeovy multiplikátory a je bez omezení za cenu použití dvou sad proměnných, zatímco duální formulace má často malou dimenzi za cenu použití složitějších matic. Duální problém je často lépe podmíněný. Povšimněte si, že levá nerovnost v (13.9) může být nahrazena rovností. Duální formulace je základem tzv. FETI metod rozložení oblasti pro paralelní řešení rozsáhlých úloh mechaniky.

13.2 Dualita pro kvadratické programování s nerovnostními vazbami

Zopakujeme si, s nezbytnými modifikacemi, předchozí postup pro úlohu konvexního kvadratického programování s nerovnostmi

$$\min_{x \in \Omega} f(x), \quad f(x) = \frac{1}{2}x^T Ax - b^T x, \quad \Omega = \{x \in \mathbb{R}^n : Bx \leq c\}, \quad (13.10)$$

kde A je opět symetrická pozitivně definitní a B má plnou řádkovou hodnost. Z našeho předpokladu plyne, že příslušná KKT soustava

$$\nabla_x L_0(x, \lambda) = Ax - b + B^T \lambda = o, \quad (13.11)$$

$$\nabla_\lambda L_0(x, \lambda) = Bx - c \leq o, \quad (13.12)$$

$$\lambda \geq o, \quad (13.13)$$

$$\lambda^T (Bx - c) = 0 \quad (13.14)$$

má jediné řešení $(\bar{x}, \bar{\lambda})$. Vyjádřeme si opět x z (13.11) a následně dosadíme za x do (13.12). Rovnici (13.11) můžeme opět považovat za podmínku, jejíž splnění zaručuje, že

$$L_0(x, \lambda) = \frac{1}{2}x^T Ax - b^T x + \lambda^T (Bx - c)$$

dosahuje v x minimum vzhledem k x . Pro dané $\lambda \in \mathbb{R}^m$ je tedy *první krok opět ekvivalentní nalezení argumentu minima*

$$x = x(\lambda) = A^{-1}(b - B^T \lambda)$$

Lagrangiánu $L_0(x, \lambda)$ vzhledem k x . Pomocí tohoto pozorování si můžeme znovu vyjádřit explicitně *duální funkci*

$$\Theta(\lambda) = -\frac{1}{2}\lambda^T BA^{-1}B^T \lambda + (BA^{-1}b - c)^T \lambda - \frac{1}{2}b^T A^{-1}b$$

a její gradient

$$\nabla \Theta(\lambda) = -BA^{-1}B^T \lambda + (BA^{-1}b - c). \quad (13.15)$$

Dosadíme-li si nyní za $x = x(\lambda)$ do (13.12), dostaneme nerovnost

$$-BA^{-1}B^T \lambda + (BA^{-1}b - c) \leq o.$$

Porovnáme-li levou stranu poslední nerovnosti s explicitním výrazem (13.15) pro $\nabla \Theta(\lambda)$, můžeme si poslední nerovnost zapsat ve tvaru

$$\nabla \Theta(\lambda) \leq o,$$

což je totéž jako

$$\nabla \Theta(\lambda) + I\mu = o, \quad \mu \geq o, \quad \mu^T \nabla \Theta(\lambda) = 0$$

nebo

$$-\nabla \Theta(\lambda) - I\mu = o, \quad \mu \geq o, \quad \mu^T \nabla \Theta(\lambda) = 0. \quad (13.16)$$

Z (13.14) pak dostaneme podmínku komplementarity

$$\lambda^T \nabla \Theta(\lambda) = 0. \quad (13.17)$$

Jelikož matice $BA^{-1}B^T$, což je Hessián funkce $-\Theta$, je podle předpokladu pozitivně definitní, dostáváme tak, že (13.16) je ekvivalentní KKT podmínce pro argument minima $-\Theta$ za podmínky $-\lambda \leq o$, což je současně argument maxima Θ vzhledem k $\lambda \geq o$. Proto KKT dvojice $(\bar{x}, \bar{\lambda})$ pro problém (13.10) s nerovnostními vazbami řeší též *min-max problém*

$$L_0(\bar{x}, \bar{\lambda}) = \max_{\substack{\lambda \in \mathbb{R}^m \\ \lambda \geq 0}} \min_{x \in \mathbb{R}^n} L_0(x, \lambda), \quad (13.18)$$

$\bar{\lambda}$ řeší *duální problém*

$$\Theta(\bar{\lambda}) = \max_{\substack{\lambda \in \mathbb{R}^m \\ \lambda \geq 0}} \Theta(\lambda),$$

a protože \bar{x} je přípustný vektor, který splňuje podmínku komplementarity (13.17), platí

$$f(\bar{x}) = L_0(\bar{x}, \lambda) = L_0(\bar{x}, \bar{\lambda}) = \Theta(\bar{\lambda}). \quad (13.19)$$

Navíc zřejmě platí

$$f(\bar{x}) = L_0(\bar{x}, \bar{\lambda}) = \min_{x \in \mathbb{R}^n} L_0(x, \bar{\lambda}) \leq L_0(x, \bar{\lambda}), \quad x \in \mathbb{R}^n. \quad (13.20)$$

Existuje také ještě jeden ekvivalentní problém související s penalizační metodou. Jelikož

$$\sup_{\substack{\lambda \in \mathbb{R}^m \\ \lambda \geq 0}} L_0(x, \lambda) = \infty \text{ pro } x \notin \Omega \quad \text{a} \quad \sup_{\substack{\lambda \in \mathbb{R}^m \\ \lambda \geq 0}} L_0(x, \lambda) = f(x) \text{ pro } x \in \Omega,$$

platí, že řešení \bar{x} KKT soustavy (13.11)-(13.14) splňuje

$$f(\bar{x}) = \min_{x \in \mathbb{R}^n} \sup_{\substack{\lambda \in \mathbb{R}^m \\ \lambda \geq 0}} L_0(x, \lambda). \quad (13.21)$$

Porovnáním (13.21) s (13.18) a (13.19) dostaneme *relaci duality*

$$\max_{\substack{\lambda \in \mathbb{R}^m \\ \lambda \geq 0}} \min_{x \in \mathbb{R}^n} L_0(x, \lambda) = \min_{x \in \mathbb{R}^n} \sup_{\substack{\lambda \in \mathbb{R}^m \\ \lambda \geq 0}} L_0(x, \lambda).$$

Použijeme-li (13.19) a (13.20), dostaneme, že $(\bar{x}, \bar{\lambda})$ řeší sedlobodový problém najít $(\bar{x}, \bar{\lambda})$ tak, aby pro každé $x \in \mathbb{R}^n$ a $\lambda \in \mathbb{R}^m$ platilo

$$L_0(\bar{x}, \lambda) \leq L_0(\bar{x}, \bar{\lambda}) \leq L_0(x, \bar{\lambda}).$$

Odvodili jsme si tak dva problémy *bez omezení*, které jsou ekvivalentní původnímu problému s nerovnostními omezeními (13.10). Sedlobodová formulace obsahuje explicitně Lagrangeovy multiplikátory a je bez omezení za cenu použití dvou sad proměnných, zatímco duální formulace má často malou dimenzi s podmínkou nezápornosti za cenu použití složitějších matic. Nahrazení obecnějších nerovností podmínkou nezápornosti je velkým přínosem pro numerické řešení, neboť je v tomto případě velmi snadné implementovat eukleidovskou projekci na přípustnou množinu. Duální problém je často lépe podmíněný. Duální formulace je základem efektivních FETI metod rozložení oblasti pro paralelní řešení rozsáhlých kontaktních úloh mechaniky.

13.3 Vizualizační lemma

Pojmy a základní myšlenky spojené s dualitou pro obecnější konvexní problém

$$\min_{h(x) \leq 0} f(x), \quad (13.22)$$

který je zadaný pomocí dostatečně hladkých funkcí $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ a má řešení $(\bar{x}, \bar{\lambda})$, si můžeme znázornit pomocí nadrovin v \mathbb{R}^{m+1} a zobrazení

$$\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^{m+1},$$

které je definováno předpisem

$$\Phi(x) = (f(x), h(x)).$$

Budeme přitom sledovat výklad Bertsekase [1]. Pro náš účel je *nadrovina* H přirozeně zadaná normálovým vektorem $n \in \mathbb{R}^{m+1}$, $n = (n_f, n_h)$, $n_f \in \mathbb{R}$, $n_h \in \mathbb{R}^m$ a konstantou $c \in \mathbb{R}$ pomocí

$$H = \{(y_f, y_h) \in \mathbb{R}^{m+1} : n_f^T y_f + n_h^T y_h = c\}.$$

Známe-li tedy nějaký bod $\bar{y} = (\bar{y}_f, \bar{y}_h) \in \mathbb{R}^{m+1}$, potom

$$H = \{(y_f, y_h) \in \mathbb{R}^{m+1} : n_f^T y_f + n_h^T y_h = n_f^T \bar{y}_f + n_h^T \bar{y}_h\}.$$

Tyto nadroviny definují dva poloprostory, a to *kladný poloprostor*

$$H^+ = \{(y_f, y_h) \in \mathbb{R}^{m+1} : n_f^T y_f + n_h^T y_h \geq n_f^T \bar{y}_f + n_h^T \bar{y}_h\}$$

a *záporný poloprostor*

$$H^- = \{(y_f, y_h) \in \mathbb{R}^{m+1} : n_f^T y_f + n_h^T y_h \leq n_f^T \bar{y}_f + n_h^T \bar{y}_h\}.$$

Nadroviny s normálou $n = (n_f, n_h)$, pro kterou platí $n_f = 0$ se nazývá vertikální, v opačném případě, tedy když $n_f \neq 0$, se n nazývá nevertikální. Nevertikální normála se nazývá normalizovaná, jestliže $n_f = 1$. Tyto definice nám umožňují spojit teorii Lagrangeových multiplikátorů s geometrií oboru hodnot Φ v následujícím lemmatu, které je ilustrováno na obr. 13.1.

Lemma 13.1. (Vizualizační lemma)

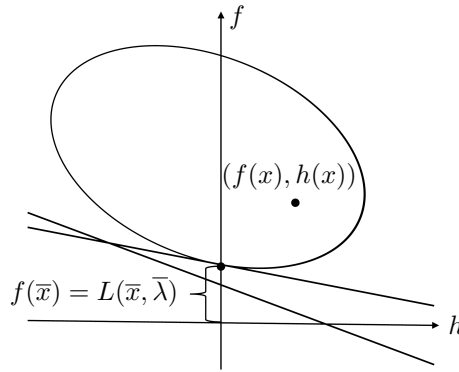
(i) Nadrovina s normálou $(1, \lambda)$, ve které leží bod $(f(x), h(x))$, protíná svislou osu $\{(z, o) : z \in \mathbb{R}\}$ v

$$z = L(f(x), h(x)).$$

(ii) Nejvyššího průsečíku nadroviny $H = \{(y_f, y_h) \in \mathbb{R}^{n+1} : 1y_f + \lambda^T y_h = c\}$, která splňuje $H^+ \supseteq \mathcal{H}(\Phi)$, se svislou osou $\{(z, o) : z \in \mathbb{R}\}$ je dosaženo v

$$z = \inf_{x \in \mathbb{R}^n} L(x, \lambda).$$

(iii) $\bar{\lambda}$ je Lagrangeův multiplikátor pro řešení konvexní úlohy (13.22), právě když $\bar{\lambda} \geq o$ a nejvyššího průsečíku nadroviny $H = \{(y_f, y_h) \in \mathbb{R}^{n+1} : 1y_f + \lambda^T y_h = c\}$, která splňuje $H^+ \supseteq \mathcal{H}(\Phi)$, se svislou osou $\{(z, o) : z \in \mathbb{R}\}$, je dosaženo v $y_f = f(\bar{x})$.



Obr. 13.1: Ilustrace vizualizačního lematu

Důkaz. (i) Nadrovina určená normálou $(1, \lambda)$ a bodem $(f(x), h(x))$ je množina všech (y_f, y_h) , které splňují

$$y_f + \lambda^T y_h = f(x) + \lambda^T h(x) = L(x, \lambda).$$

Jediný vektor na svislé ose $\{(z, o) : z \in \mathbb{R}\}$ je $(L(x, \lambda), o)$.

(ii) Nadrovina s normálou $(1, \lambda)$, která protíná vertikální osu ve výšce c je množina vektorů, které splňují rovnici

$$y_f + \lambda y_h = c,$$

a tato nadrovina obsahuje množinu $\mathcal{S} = \mathcal{H}(\Phi)$ ve svém kladném poloprostoru, právě když

$$L(x, \lambda) = f(x) + \lambda^T h(x) \geq c.$$

(iii) Plyne z definice Lagrangeova multiplikátoru a části (iii). □

Kapitola 14

Lineární programování

Lineární programování se zabývá minimalizací cenového funkcionálu na polyedru, tj. na průniku poloprostorů. Je to specifická optimalizační úloha, která sice patří mezi úlohy spojité optimalizace, avšak lze ji řešit i metodami, které jsou blízké diskrétní optimalizaci. V této kapitole se seznámíme s různými formulacemi úlohy lineárního programování, s fakty z geometrie polyedru, a se dvěma zcela odlišnými přístupy k řešení.

14.1 Formulace úlohy lineárního programování

Budeme se zabývat řešením úlohy najít

$$\min_{x \in \Omega} b^T x, \quad \Omega = \{x \in \mathbb{R}^n : B_{\mathcal{E}}x = c_{\mathcal{E}} \quad \text{a} \quad B_{\mathcal{I}}x \leq c_{\mathcal{I}}\}, \quad (14.1)$$

kde $b \in \mathbb{R}^n$, $c \in \mathbb{R}^m$, $B \in \mathbb{R}^{m \times n}$, a \mathcal{E}, \mathcal{I} je rozklad množiny indexů $\{1, \dots, m\}$.

Formulace úlohy (14.1) je sice obecná, avšak není vhodná k řešení. Ukazuje se, že vhodnější je nahradit obecné lineární nerovnosti nezáporným omezením pomocí nové proměnné

$$y = c_{\mathcal{I}} - B_{\mathcal{I}}x,$$

takže $x \in \Omega$, právě když

$$B_{\mathcal{E}}x = c_{\mathcal{E}}, \quad B_{\mathcal{I}}x + y = c_{\mathcal{I}} \quad \text{a} \quad y \geq 0.$$

Pro některé algoritmy je dále výhodné, když jsou všechny proměnné nezáporné. Toho dosáhneme snadno, když položíme

$$x = z - w, \quad z \geq 0 \quad \text{a} \quad w \geq 0.$$

Označíme-li si

$$A = \begin{bmatrix} B_{\mathcal{E}} & -B_{\mathcal{E}} & O \\ B_{\mathcal{I}} & -B_{\mathcal{I}} & I \end{bmatrix}, \quad \bar{x} = \begin{bmatrix} z \\ w \\ y \end{bmatrix}, \quad \bar{b} = \begin{bmatrix} b \\ -b \\ o \end{bmatrix},$$

můžeme zapsat úlohu lineárního programování ve standardním tvaru

$$\min_{\bar{x} \in \bar{\Omega}} \bar{b}^T \bar{x}, \quad \bar{\Omega} = \{\bar{x} \in \mathbb{R}^{n+|\mathcal{I}|} : A\bar{x} = c \quad a \quad \bar{x} \geq o\}. \quad (14.2)$$

14.2 Podmínky minima a dualita

Napíšeme-li si podmínky minima pro úlohu minima ve standardním tvaru

$$\min_{x \in \Omega} b^T x, \quad \Omega = \{x \in \mathbb{R}^n : Ax = c \quad a \quad x \geq o\}, \quad (14.3)$$

dostaneme

$$b - A^T \lambda - \mu = o, \quad \mu \geq o, \quad Ax = c, \quad x \geq o, \quad x^T \mu = 0.$$

Lagrangián pro úlohu (14.3) má tvar

$$L(x, \lambda, \mu) = b^T x + \lambda^T (c - Ax) - \mu^T x,$$

takže si můžeme odvodit duální funkci pro $\lambda \in \mathbb{R}^n$ a $\mu \geq o$ ve tvaru

$$\begin{aligned} \Theta(\lambda, \mu) &= \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu) = \inf_{x \in \mathbb{R}^n} b^T x + \lambda^T (c - Ax) - \mu^T x \\ &= \inf_{x \in \mathbb{R}^n} x^T (b - A^T \lambda) + \lambda^T c - \mu^T x \\ &= \begin{cases} \lambda^T c & \text{pokud } b - A^T \lambda \geq o, \quad x \geq o, \\ -\infty & \text{pokud existuje některé } i \in \mathcal{E}: [b - A^T \lambda]_i < 0, \quad x_i < 0. \end{cases} \end{aligned}$$

Duální problém je tedy opět lineární problém, jehož cílem je najít $\lambda \in \mathbb{R}^m$, které řeší

$$\lambda \in \mathbb{R}^m, \quad \max_{\mu \geq o} \Theta(\lambda, \mu) = \max_{\lambda \in \mathbb{R}^m} \lambda^T c \quad \text{pro } A^T \lambda \leq b.$$

Jak již víme, složky λ a $\mu (= b - A^T \lambda)$ popisují citlivost cenové funkce na narušení vazeb. V ekonomických aplikacích je lze často interpretovat jako stínové ceny. Duální formulace je užitečná pro odvození některých efektivních algoritmů pro řešení úloh lineárního programování.

14.3 Metoda vnitřního bodu

Nyní si stručně ukážeme, jak adaptovat metodu vnitřního bodu na řešení úlohy lineárního programování ve standardním tvaru

$$\min_{x \in \Omega} b^T x, \quad \Omega = \{x \in \mathbb{R}^n : Ax = c \quad a \quad x \geq o\}. \quad (14.4)$$

Budeme přitom předpokládat, že řádky vazební matice A jsou lineárně nezávislé, a zapíšeme si duální úlohu ve tvaru

$$\max_{\lambda \in \mathbb{R}^m} \lambda^T c \quad \text{pro} \quad A^T \lambda + \mu = b, \quad \mu \geq o.$$

Podmínky minima pro úlohu (14.4) si zapíšeme ve standardním tvaru

$$\begin{aligned} A^T \lambda + \mu &= b, \\ Ax &= c, \\ x^T \mu &= 0, \\ x &\geq o, \\ \mu &\geq o. \end{aligned}$$

K řešení této soustavy použijeme variantu Newtonovy metody, a to tak, že budeme v každém kroku zajišťovat splnění podmínky nezápornosti x a μ . Povšimněte si, že ostatní relace jsou buď lineární nebo mírně nelineární.

Přepíšeme si nejprve podmínku komplementarity ve tvaru

$$XSe = o, \quad X = \text{diag}[x_1, \dots, x_n]^1, \quad S = \text{diag}[\mu_1, \dots, \mu_n], \quad e = [1, \dots, 1]^T$$

a definujme si zobrazení $F : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$ pomocí

$$F(x, \lambda, \mu) = \begin{bmatrix} A^T \lambda + \mu - b \\ Ax - c \\ XSe \end{bmatrix}.$$

Snadno si ověříme, že Jacobián F má tvar

$$J(x, \lambda, \mu) = \begin{bmatrix} O & A^T & I \\ A & O & O \\ S & O & X \end{bmatrix},$$

takže Newtonův směr je definován rovnicí

$$F(x, \lambda, \mu) + J(x, \lambda, \mu) \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta \mu \end{bmatrix} = o.$$

¹Symbol *diag* označuje čtvercovou matici, která obsahuje na hlavní diagonále prvky zapsané ve vektoru, který je uveden za tímto symbolem. Jiné prvky jsou nulové.

Je-li (x^k, λ^k, μ^k) přípustná vnitřní iterace, tj. $x^k > 0$, $\mu^k > 0$, pak definujeme novou iteraci pomocí

$$(x^{k+1}, \lambda^{k+1}, \mu^{k+1}) = (x^k, \lambda^k, \mu^k) + \alpha(\Delta x^k, \Delta \lambda^k, \Delta \mu^k),$$

kde $\alpha \in (0, 1)$ zvolíme tak, aby nová iterace byla opět vnitřní iterace.

Ukazuje se, že přirozený požadavek na splnění podmínky komplementarity vede k příliš krátkému kroku. Lepší výsledek se obvykle dosáhne, když se poslední podmínka částečně uvolní. Proveďte se to tak, že si zavedeme průměrnou míru narušení podmínky komplementarity

$$\delta = \frac{x^T \mu}{n}$$

a vhodný centralizující parametr $\sigma_k > 0$, s jejichž pomocí zapíšeme relaxovanou rovnici pro Newtonův směr ve tvaru

$$\begin{bmatrix} O & A^T & I \\ A & O & O \\ S^k & O & X^k \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta \lambda^k \\ \Delta \mu^k \end{bmatrix} = \begin{bmatrix} b - A^T \lambda^k - \mu^k \\ c - A x^k \\ -X^k S^k e + \sigma_k \delta_k e \end{bmatrix}.$$

Konvergenci k řešení zajistíme zmenšováním σ_k .

14.4 Geometrie simplexové metody

V sekci 14.3 jsme si ukázali, že úlohy lineárního řešení je možné řešit metodami, které byly navrženy pro řešení obecnějších nelineárních úloh. Úloha lineárního programování má však specifickou strukturu, kterou je možno využít pro navržení zcela jiného algoritmu. Hlavní myšlenka simplexové metody, která využívá tuto specifickou strukturu, je patrná z obr. 14.1.

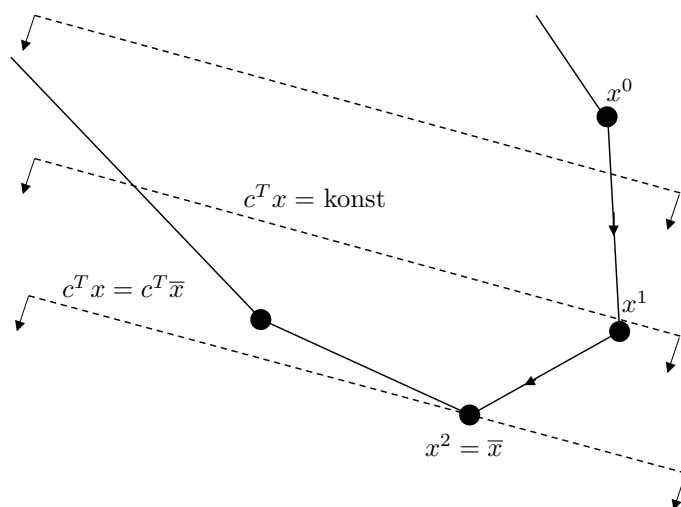
Podle obr. 14.1 se dá především očekávat, že minimum nemůže být dosaženo ve vnitřku přípustné množiny Ω , ale někde na hranici. Toto pozorování si můžeme snadno dokázat, když si uvědomíme, že přípustná množina Ω je konvexní tj. pro libovolné $\alpha \in (0, 1)$ a $y, z \in \Omega$ platí

$$\alpha y + (1 - \alpha)z \in \Omega,$$

a lineární cenová funkce $f(x) = c^T x$ je konkávní, tedy splňuje

$$f(\alpha y + (1 - \alpha)z) \geq \alpha f(y) + (1 - \alpha)f(z) \geq \min\{f(y), f(z)\}.$$

Obrázek však naznačuje nejen to, že při hledání minima můžeme vynechat vnitřní body Ω , ale že minimum je dosaženo v bodech lomu hranice, což jsou body konvexní hranice Ω . Toto tvrzení lze dokázat (za předpokladu, že přípustná množina Ω je neprázdná) i formálně, a to i v případě, že Ω je větší než konvexní obal bodů konvexní hranice, tak jako v našem případě. My si ho dokážeme v příští sekci.



Obr. 14.1: Geometrické znázornění simplexové metody

Postup řešení úlohy (14.3) simplexovou metodou spočívá v systematickém procházení bodů konvexní hranice, počínaje vrcholem x^0 . Jako další iteraci zvolíme jeden ze sousedních bodů hranice, obvykle ve směru nejbližším zápornému gradientu cenové funkce, tedy $-c$. Hlavním problémem je algebraicky popsat body konvexní hranice a pravidlo nalezení další iterace, tedy přechodu od vrcholu k vrcholu s menší hodnotou cenové funkce, což si ukážeme v další sekci.

14.5 Konvexní hranice simplexu

Obsahem následující věty je algebraická reprezentace bodů konvexní hranice přípustné množiny Ω .

Věta 14.1. *Necht*

$$\Omega = \{x \in \mathbb{R}^n : Ax = c \text{ a } x \geq o\} \neq \emptyset$$

a A je $m \times n$ matice plné hodnosti, $m < n$, jejíž sloupce lze uspořádat tak, že

$$A = [B, N],$$

kde B je regulární matice řádu m . Potom x je bodem konvexní hranice Ω , právě když

$$x = \begin{bmatrix} x_B \\ x_N \end{bmatrix} = \begin{bmatrix} B^{-1}c \\ o \end{bmatrix}$$

Důkaz. Ukažme si nejprve, že x nelze zapsat jako konvexní kombinaci žádných bodů $y, z \in \Omega$, $y \neq z$, tj. z $x = \alpha y + (1 - \alpha)z$, $\alpha \in (0, 1)$ plyne $\alpha = 0$ nebo $\alpha = 1$. Skutečně, z $y \geq o$, $z \geq o$ a předpokladu plyne bezprostředně, že $y_N = z_N = o$, takže

$$Ax = Bx_B = By_B = Bz_B = c,$$

tedy $y_B = z_B = B^{-1}c$ a $x = y = z$, což je spor.

Obráceně, předpokládejme, že x je bod konvexní hranice. Bez újmy na obecnosti můžeme předpokládat, že x má dvě komponenty, nenulovou $x_B \in \mathbb{R}^p$ a nulovou x_N . Ukážeme si sporem, že sloupce a_1, \dots, a_p matice A odpovídající nenulovým komponentám jsou nezávislé. Předpokládejme tedy, že jsou lineárně závislé, takže existují $\alpha_1, \dots, \alpha_p$ tak, že nejsou všechny nulové a

$$\alpha_1 a_1 + \dots + \alpha_p a_p = o.$$

Pro vektor α dimenze n , jehož nenulové složky jsou právě $\alpha_1, \dots, \alpha_p$, tedy platí

$$A\alpha = o \text{ a } \alpha \neq o.$$

Jelikož $x_B > o$, bude pro dostatečně malé nenulové $\bar{\alpha}$ platit

$$x + \bar{\alpha}\alpha \in \Omega \text{ a } x - \bar{\alpha}\alpha \in \Omega,$$

takže x lze zapsat ve sporu s předpokladem jako konvexní kombinaci dvou různých přípustných vektorů ve tvaru

$$x = \frac{1}{2}(x + \bar{\alpha}\alpha) + \frac{1}{2}(x - \bar{\alpha}\alpha).$$

Tím jsme dokázali, že sloupce a_1, \dots, a_p jsou nezávislé, z čehož také vyplývá, že $p \leq m$. Pokud $m = p$, označme si $B = [a_1, \dots, a_m]$, $N = [a_{m+1}, \dots, a_n]$, takže $A = [B, N]$, a označme si x_B a x_N vektory odpovídajících složek x . Pak $x_N = o$ a

$$Ax = Bx_B = c,$$

tedy $x_B = B^{-1}c$, což je důkaz tvrzení pro $m = p$. Pokud $p < m$, doplníme a_1, \dots, a_p o další sloupce do báze oboru hodnot matice A (tedy \mathbb{R}^m) a přeuspořádáme x, A tak, že doplňující vektory vzájemně zaměníme. \square

Jak jsme se zmínili výše, simplexový algoritmus je založen na tvrzení, že minimum je dosaženo v bodech konvexní hranice. Přírozený důkaz je založen na konkávnosti cenové funkce a na reprezentaci Ω pomocí konvexní kombinace bodů konvexní hranice a nezáporné kombinace tzv. extrémních směrů. Důkaz potřebné věty o reprezentaci je však za hranicí tohoto úvodního textu, uvedeme si proto jiný důkaz, který je podobný důkazu věty 14.1.

Věta 14.2. *Nechť úloha (14.2) má řešení. Potom existuje řešení \bar{x} , které je bodem konvexní hranice přípustné množiny Ω .*

Důkaz. Nechť \bar{x} je řešení s nejmenším počtem nenulových komponent p . Bez újmy na obecnosti můžeme předpokládat, že \bar{x} má dvě komponenty, nenulovou $\bar{x}_B \in \mathbb{R}^p$ a nulovou \bar{x}_N .

Dokážeme si nejprve sporem, že sloupce matice A odpovídající nenulovým komponentám \bar{x} jsou nezávislé, takže existují $\alpha_1, \dots, \alpha_p$ tak, že nejsou všechny nulové a

$$\alpha_1 a_1 + \dots + \alpha_p a_p = o.$$

Pro vektor α dimenze n , jehož nenulové složky jsou právě $\alpha_1, \dots, \alpha_p$, tedy platí

$$A\alpha = o \quad \text{a} \quad \alpha \neq o.$$

Jelikož $\bar{x}_B > o$, bude pro dostatečně malé nenulové $\bar{\alpha}$ platit

$$\bar{x} + \bar{\alpha}\alpha \in \Omega \quad \text{a} \quad \bar{x} - \bar{\alpha}\alpha \in \Omega.$$

Pro každé $x \in \Omega$ však platí

$$f(\bar{x}) = b^T \bar{x} \leq b^T x = f(x),$$

takže pro dostatečně malé $\bar{\alpha}$ dostaneme

$$b^T \bar{x} \leq b^T (\bar{x} + \bar{\alpha}\alpha) \quad \text{a} \quad b^T \bar{x} \leq b^T (\bar{x} - \bar{\alpha}\alpha),$$

tedy pro jakékoliv $\bar{\alpha}$, které je dostatečně malé v absolutní hodnotě, platí

$$b^T \bar{x} = b^T (\bar{x} + \bar{\alpha}\alpha).$$

Pro vhodnou volbu $\bar{\alpha}$ tedy můžeme dosáhnout toho, že vektor

$$\bar{x}_{\bar{\alpha}} = \bar{x} + \bar{\alpha}\alpha$$

má $p + 1$ nulových složek, což je spor s předpokladem, že \bar{x} je řešení s nejmenším počtem nenulových komponent.

Dokázali jsme tedy, že sloupce matice A s indexy nenulových složek vektoru \bar{x} jsou nezávislé. Pokud $p < m$, doplníme a_1, \dots, a_p o další sloupce do báze oboru hodnot matice A (tedy \mathbb{R}^m). Po vhodném přeuspořádání \bar{x} , A bude mít vektor \bar{x} složky s indexy $m + 1, \dots, n$ nulové a prvních m sloupců A bude nezávislých. Podle věty 14.1 je tedy \bar{x} bodem konvexní hranice. \square

14.6 Krok simplexové metody

Nejprve využijeme větu 14.2 k zápisu podmínky, kterou splňuje řešení úlohy (14.2), ve vhodnější formě než jsou standardní KKT podmínky.

Nechť x je bod konvexní hranice, jehož prvních m komponent je nenulových, takže podle věty 14.2 platí

$$x = \begin{bmatrix} x_B \\ x_N \end{bmatrix} = \begin{bmatrix} B^{-1}c \\ o \end{bmatrix}.$$

Pro libovolný bod $y \in \Omega$ s vektorovými složkami y_B, y_N platí

$$Ay = By_B + Ny_N = c,$$

takže

$$y_B = B^{-1}c - B^{-1}Ny_N$$

a pro cenovou funkci určenou vektorem b s vektorovými složkami b_B, b_N dostaneme

$$\begin{aligned} b^T y &= b_B^T y_B + b_N^T y_N = b_B^T B^{-1}c + (b_N^T - b_B^T B^{-1}N)y_N \\ &= b^T x + (b_N^T - b_B^T B^{-1}N)y_N. \end{aligned}$$

Jelikož $y_N \geq o$, je x řešení, právě když

$$b_N^T - b_B^T B^{-1}N \geq o.$$

Jestliže tedy x není řešením, pak existuje $j > m$ tak, že platí

$$b_j - b_B^T B^{-1} a_j < 0.$$

Budeme proto hledat bod y konvexní hranice Ω , který sousedí s x tak, aby $y_j > 0$. Budeme ho hledat ve tvaru

$$y = x + \lambda d, \quad \lambda > 0, \quad d_N = e_{j-m},$$

kde e je sloupec jednotkové matice řádu $n - m$. Má-li platit $y \in \Omega$, pak musí platit

$$Ay = Ax + \lambda(Bd_B + Nd_N) = Ax + \lambda(Bd_B + a_j) = c.$$

Jelikož $Ax = c$, dostáváme tak

$$Bd_B + a_j = 0, \quad \text{tj.} \quad d = \begin{bmatrix} d_B \\ d_N \end{bmatrix} = \begin{bmatrix} -B^{-1}a_j \\ e_{j-m} \end{bmatrix},$$

Snadno si ověříme, že pro cenovou funkci v y platí

$$f(y) = b^T y = b^T x + \lambda(b_j - b_B^T B^{-1} a_j).$$

Pro $\lambda > 0$ tedy bude platit $f(y) < f(x)$, neboť analyzujeme případ $b_j - b_B^T B^{-1} a_j < 0$.

Pro další postup budeme rozlišovat dva případy.

i. Ve směru d je možný neomezený pokles Necht $d_B = -B^{-1}a_j \geq 0$. V tomto případě $x + \lambda d \in \Omega$ pro každé $\lambda > 0$ a

$$\lim_{\lambda \rightarrow \infty} b^T x + \lambda(b_j - b_B^T B^{-1} a_j) = -\infty$$

neboť předpokládáme $b_j - b_B^T B^{-1} a_j < 0$. Tento případ nemůže nastat když úloha (14.2) má řešení.

ii. Ve směru d je bod konvexní hranice s menší hodnotou cenové funkce Necht existuje $i \in \{1, \dots, m\}$ tak, že $d_i < 0$. V tomto případě položíme

$$\bar{\lambda} = \max \{ \lambda > 0 : x + \lambda d \in \Omega \}.$$

Snadno se ověří, že

$$y = x + \bar{\lambda} d$$

je bod konvexní hranice, pro který platí $f(x) < f(y)$.

Tak jsme si ukázali, že pokud máme bod konvexní hranice s m kladnými složkami (takové body nazýváme také *regulární*), je to buď bod, ve kterém je dosaženo minima, nebo můžeme najít jiný bod konvexní hranice s menší hodnotou cenové funkce.

Postup lze modifikovat (i když to není zcela triviální) i pro body konvexní hranice, které nejsou regulární. Jelikož bodů konvexní hranice je konečný počet a žádný bod se nevygeneruje dvakrát (cenová funkce klesá v každé iteraci), najde se uvedeným algoritmem řešení problému v konečném počtu kroků.

Poznamenejme, že předchozí úvaha nezaručuje, že se pro větší úlohy řešení opravdu najde. Důvodem je počet bodů konvexní hranice, který se rovná

$$p = \binom{n}{m}.$$

Jsou dokonce známé problémy, pro které je opravdu možné najít řešení až po p krocích. Simplexový algoritmus má proto exponenciální složitost. Ve většině případů je však k řešení třeba $2m$ až $3m$ kroků. Nalezení algoritmu s polynomiální složitostí byla událost, kterou zaznamenal i Time magazín.

Kapitola 15

Úvod do metod nehladké optimalizace

V této kapitole se zaměříme na řešení úlohy

$$\min_{x \in \mathbb{R}^n} f(x), \quad (15.1)$$

kde funkce f je pouze spojitá, a není tedy zaručena její diferencovatelnost. K tomuto účelu se nejprve budeme podrobně věnovat základním nástrojům potřebným pro práci s funkcemi, které nejsou spojitě diferencovatelné. Podíváme se na rozšíření klasického diferenciálního počtu pro spojitě diferencovatelné (hladké) funkce. Toto rozšíření nám přináší *Clarkeův kalkul* zavádějící zobecněný gradient i v bodech funkce, kde uvažovaná funkce není spojitě diferencovatelná. Clarkeův kalkul je pro své vlastnosti vhodný k algoritmizaci hledání minima funkcí, jež nejsou spojitě diferencovatelné. Jiným kalkulem, který je vhodný pro práci s nehladkými funkcemi je *Morduchovičův kalkul*. Ten nám umožňuje „přísnější“ ověření podmínek optimality. Tento kalkul jde však nad rámec tohoto textu a proto se mu nadále věnovat nebudeme. Mnoho informací o těchto kalkulech lze najít v [3] a [13]. Nakonec se podíváme na možnosti řešení některých nehladkých problémů. Buď nepřímo převedením na hladkou úlohu anebo přímo použitím některého z algoritmů vhodného pro nehladkou optimalizaci.

15.1 Základní pojmy

Nyní se seznámíme s některými základními pojmy z oblasti nehladké analýzy.

Definice 15.1. Necht množina $M \subset \mathbb{R}^n$, funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a K je nezáporné reálné číslo:

(i) Necht pro funkci f je splněna podmínka

$$|f(x') - f(x'')| \leq K \|x' - x''\| \quad \forall x', x'' \in M.$$

Pak řekneme, že funkce f je *lipschitzovsky spojitá* (s modulem K) na množině M .

(ii) Necht pro funkci f je splněna podmínka

$$\exists \varepsilon > 0 : |f(x') - f(x'')| \leq K \|x' - x''\| \quad \forall x', x'' \in x + \varepsilon B,$$

kde B je jednotková koule. Pak řekneme, že funkce f je *lipschitzovsky spojitá v okolí bodu* $x \in \mathbb{R}^n$ (s modulem K).

(iii) Necht funkce f je lipschitzovsky spojitá v okolí každého $x \in \mathbb{R}^n$. Pak řekneme, že funkce f je *lokálně lipschitzovsky spojitá* na \mathbb{R}^n .

Je zřejmé, že lipschitzovská spojitost na množině je silnější podmínka než lokální lipschitzovská spojitost. Např. funkce $f(x) = x^2$ je na \mathbb{R} lokálně lipschitzovská spojitá, ale není lipschitzovsky spojitá.

Následující tvrzení nám dává informaci o „velikosti“ podmnožiny definičního oboru, ve které funkce, která je lokálně lipschitzovsky spojitá funkce na \mathbb{R}^n , není diferencovatelná.

Věta 15.2. (Rademacher) Necht funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}$ je lokálně lipschitzovsky spojitá na \mathbb{R}^n a necht $D = \{x \in \mathbb{R}^n \mid f \text{ je diferencovatelná v } x\}$. Potom Lebesguova míra μ množiny bodů, ve kterých funkce f není diferencovatelná, tj. množiny $(\mathbb{R}^n \setminus D)$, je nula.

Z předchozího tvrzení plyne, že v případě lokálně lipschitzovsky spojitě funkce máme zaručenu její diferencovatelnost „skoro všude“. V dalším textu se budeme věnovat řešení úlohy (15.1), kde funkce f je lokálně lipschitzovsky spojitá. Nyní si zavedeme základní pojmy Clarkeova kalkulu.

Definice 15.3. Necht funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}$ je lipschitzovsky spojitá v okolí bodu x a necht v je vektor v \mathbb{R}^n . Clarkeova zobecněná směřová derivace funkce f v bodě x ve směru v , kterou značíme $f^0(x; v)$ je definována předpisem

$$f^0(x; v) = \limsup_{y \rightarrow x, t \downarrow 0} \frac{f(y + tv) - f(y)}{t},$$

kde y je vektor v \mathbb{R}^n a t je kladné reálné číslo.

Definice 15.4. Necht funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}$ je lipschitzovsky spojitá v okolí bodu x . Clarkeův zobecněný gradient funkce f v bodě x , který značíme $\partial f(x)$, je množina

$$\partial f(x) = \{\zeta \in \mathbb{R}^n \mid f^0(x; v) \geq \langle \zeta, v \rangle \forall v \in \mathbb{R}^n\}.$$

Prvky $\partial f(x)$ nazýváme *Clarkeovými subgradienty* funkce f v bodě x .

Poznámka 15.5. V některých textech - zejména v těch, které se věnují analýze konvexních nehladkých funkcí - se místo pojmu zobecněný gradient používá pojem *subdiferenciál*. V tomto textu budeme používat pojem zobecněný gradient.

Poznámka 15.6. Jestliže funkce f je spojitě diferencovatelná v bodě x , pak $f^0(x; v) = f'(x; v) = \langle \nabla f(x), v \rangle$, $\forall v \in \mathbb{R}^n$ a z toho vyplývá $\partial f(x) = \nabla f(x)$.

Vztah mezi Clarkeovou směrovou derivací a Clarkeovým zobecněným gradientem popisuje následující věta.

Věta 15.7. Necht funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}$ je lipschitzovsky spojitá v okolí bodu x . Pak platí

$$\forall v \in \mathbb{R}^n : f^0(x; v) = \max\{\langle \zeta, v \rangle : \zeta \in \partial f(x)\}.$$

Díky Rademacherovu teorému (u lokálně lipschitzovsky spojitých funkcí máme zaručenu diferencovatelnost téměř všude ve smyslu Lebesguovy míry) máme také jinou možnost, jak zkonstruovat Clarkeův zobecněný gradient. Tuto alternativní možnost nabízí níže uvedená věta.

Věta 15.8. Necht funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}$ je lipschitzovsky spojitá v okolí bodu x . Potom

$$\partial f(x) = \text{conv} \left\{ \lim_{i \rightarrow \infty} \nabla f(x_i) \mid x_i \rightarrow x, x_i \notin \Omega_f \right\},$$

kde $\Omega_f = \{x \in \mathbb{R}^n \mid f \text{ není diferencovatelná v } x\}$.

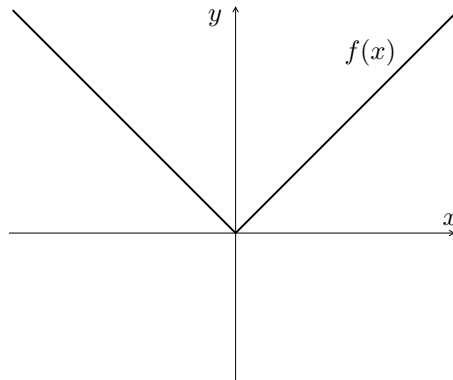
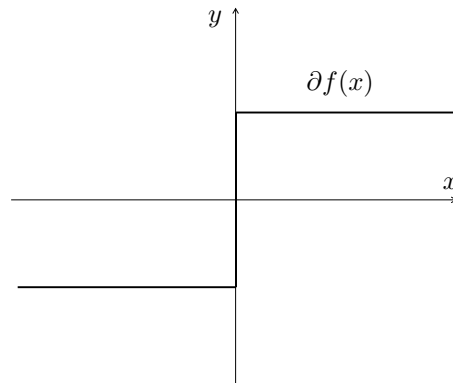
Pro ilustraci pojmu Clarkeův zobecněný gradient použijeme funkci $f(x) = |x|$ (viz obrázek 15.1).

Tato funkce je lokálně lipschitzovsky spojitá na \mathbb{R} a proto je diferencovatelná skoro všude. Jediným bodem, kde tato funkce není diferencovatelná je bod 0. Zobecněný gradient této funkce je na obrázku 15.2.

Pro hledání lokálních extrémů funkcí, které nejsou spojitě diferencovatelné, potřebujeme nástroj, který nám umožní rozpoznat, zda funkce f má v bodě x lokální extrém. To nám poskytuje následující věta.

Věta 15.9. (O lokálním extrému) Jestliže funkce f , která je lokálně lipschitzovsky spojitá na \mathbb{R}^n , nabývá v bodě x svého lokálního maxima nebo minima, potom platí

$$0 \in \partial f(x).$$

Obr. 15.1: Graf $f(x) = |x|$ Obr. 15.2: Zobecněný gradient $\partial f(x)$ funkce $f(x) = |x|$

Body splňující výše uvedenou podmínku nazýváme Clarkeovými stacionární body optimalizační úlohy

$$\min_{x \in \mathbb{R}^n} f(x),$$

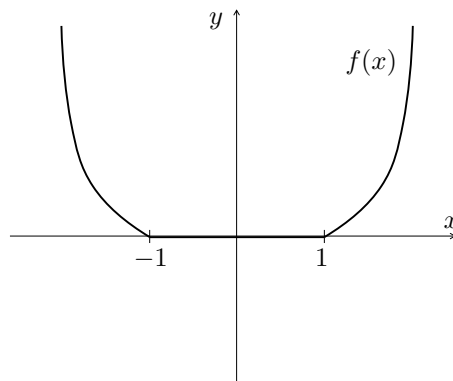
kde funkce f je lokálně lipschitzovsky spojitá na \mathbb{R}^n .

Např. funkce

$$f(x) = |x|$$

má Clarkeův zobecněný gradient v bodě 0 dán takto $\partial f(0) = \langle -1, 1 \rangle$, jak bylo ilustrováno na obrázku 15.2. Bod $x = 0$ je Clarkeův stacionární bod.

Uvedme si ještě další příklad funkce, která není diferencovatelná v celém svém definičním oboru. Takovouto funkcí je funkce $f(x) = \max \{0, \frac{1}{2}(x^2 - 1)\}$ (viz obrázek 15.3).

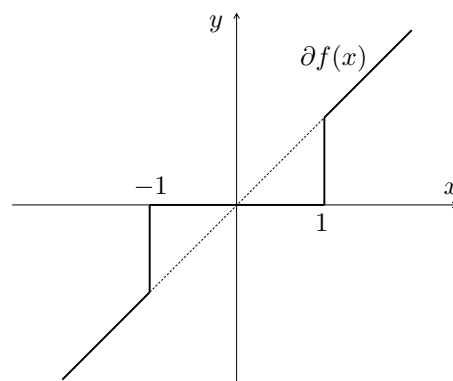
Obr. 15.3: Graf $f(x) = \max \left\{ 0, \frac{1}{2} (x^2 - 1) \right\}$

Příklad 15.10. Mějme funkci $f(x) = \max \left\{ 0, \frac{1}{2} (x^2 - 1) \right\}$. Najděte Clarkeův zobecněný gradient této funkce ve všech bodech jejího definičního oboru.

Řešení. Tato funkce je lokálně lipschitzovsky spojitá na \mathbb{R} a proto je diferencovatelná skoro všude. Jedinými body, kde tato funkce není diferencovatelná jsou body -1 a 1 . Uvedená funkce

$$f(x) = \max \left\{ 0, \frac{1}{2} (x^2 - 1) \right\}$$

má Clarkeův zobecněný gradient v intervalu $(-1, 1)$ dán takto $\partial f(x) = \nabla f(x) = 0$, v bodě -1 je její Clarkeův zobecněný gradient $\partial f(-1) = \langle -1, 0 \rangle$ a v bodě 1 je její Clarkeův zobecněný gradient $\partial f(1) = \langle 0, 1 \rangle$. Mimo interval $\langle -1, 1 \rangle$ je její Clarkeův zobecněný gradient $\partial f(x) = \nabla f(x) = x$. Vše je ilustrováno na obrázku 15.4. Body z intervalu $\langle -1, 1 \rangle$ jsou jejími Clarkeovými stacionárními body.

Obr. 15.4: Zobecněný gradient $\partial f(x)$ funkce $f(x) = \max \left\{ 0, \frac{1}{2} (x^2 - 1) \right\}$ 

Další důležitou vlastností funkce, kterou využijeme na konci této kapitoly, je polohladkost, která souvisí se vztahem mezi směrovou derivací funkce f a jejím Clarkeovým zobecněným gradientem.

Definice 15.11. O funkci $f : \mathbb{R}^n \rightarrow \mathbb{R}$ řekneme, že je *polohladká* v bodě x , jestliže funkce f je lipschitzovsky spojitá v okolí bodu x a dále platí, že pro všechny $v \in \mathbb{R}^n$ existuje limita

$$\lim_{\substack{V \in \partial f(x+tv') \\ v' \rightarrow v, t \downarrow 0}} \{Vv'\}.$$

O funkci $f : \mathbb{R}^n \rightarrow \mathbb{R}$ řekneme, že je *slabě polohladká* v bodě x , jestliže funkce f je lipschitzovsky spojitá v okolí bodu x a dále platí, že pro všechny $v \in \mathbb{R}^n$ existuje limita

$$\lim_{\substack{V \in \partial f(x+tv) \\ t \downarrow 0}} \{Vv\}.$$

Je zřejmé, že polohladkost funkce v bodě implikuje slabou polohladkost.

Pro stanovování podmínek optimality úloh omezené optimalizace je nyní potřebné zavést dva další pojmy a to *Clarkeův tečný* a *Clarkeův normálový kužel*.

Definice 15.12. Necht množina Ω je podmnožinou \mathbb{R}^n a bod $x \in \Omega$. Potom Clarkeův tečný kužel k množině Ω v bodě x , který označujeme jako $\overline{T}_\Omega(x)$, definujeme předpisem

$$\overline{T}_\Omega(x) = \{v \in \mathbb{R}^n \mid \text{dist}_\Omega^0(x; v) = 0\},$$

kde $\text{dist}_\Omega(x) = \inf_{y \in \Omega} \|x - y\|$.

Definice 15.13. Necht množina Ω je podmnožinou \mathbb{R}^n a bod $x \in \Omega$. Potom Clarkeův normálový kužel k množině Ω v bodě x , který označujeme jako $\overline{N}_\Omega(x)$, definujeme předpisem

$$\overline{N}_\Omega(x) = \{\zeta \in \mathbb{R}^n \mid \langle \zeta, v \rangle \leq 0 \forall v \in \overline{T}_\Omega(x)\}.$$

S využitím Clarkeova normálového kuželu můžeme zavést podmínku optimality pro úlohu omezené optimalizace.

Věta 15.14. *Nechť Ω je uzavřená podmnožina \mathbb{R}^n a $x \in \Omega$. Dále necht funkce $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ je lipschitzovsky spojitá v okolí bodu x a necht bod x je (vázané) lokální minimum funkce f na množině Ω . Pak platí*

$$0 \in \partial f(x) + \bar{N}_\Omega(x).$$

Body splňující výše uvedenou podmínku nazýváme Clarkeovy stacionární body optimalizační úlohy

$$\min_{x \in \Omega} f(x).$$

15.2 Nepřímé metody

Nyní se seznámíme s některými algoritmy, které jsou určeny k řešení úloh nehladké optimalizace. Nejméně propracovanou možností, jak minimalizovat funkci, která není spojitě diferencovatelná, je aproximovat tuto funkci funkcí spojitě diferencovatelnou. Toho lze dosáhnout například pomocí penalizace (viz následující ukázka). Nevýhodou tohoto postupu je jednak velmi omezené použití pouze na úzký okruh problémů a dále nutnost volit velikost penalty předem bez znalosti struktury problému.

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^m |f_i(x)|,$$

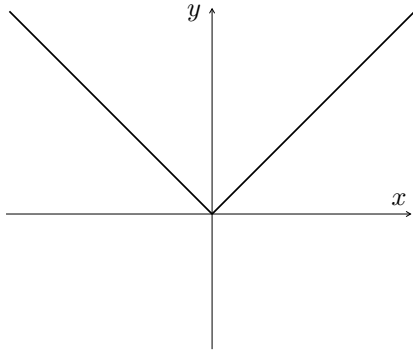
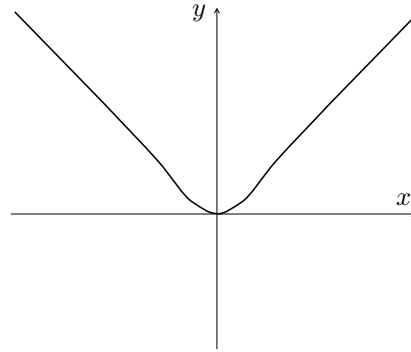
kde funkce $f_i(x)$ jsou spojitě diferencovatelné. Pomocí penalty o velikosti $\varepsilon > 0$ modifikujeme výše uvedenou úlohu na hladkou úlohu

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^m (f_i(x)^2 + \varepsilon)^{\frac{1}{2}}.$$

Tuto modifikaci můžeme předvést pro jednoduchý případ, kdy $m = 1$, $f_1(x) = x$ a penalta $\varepsilon = 0.05$. Původní funkci $f(x) = |x|$ zobrazuje obrázek 15.5, modifikovanou funkci $f(x) = (x^2 + \varepsilon)^{\frac{1}{2}}$ najdeme na obrázku 15.6.

15.3 Subgradientní metody

Nyní se již budeme věnovat metodám, které přímo pracují s nehladkými problémy. Další možností pro řešení úloh nehladké optimalizace je použití subgradientní metody. Subgradientní metoda potřebuje pro své fungování existenci procedury, která je schopna v každém bodě $x \in \mathbb{R}^n$ vypočítat hodnotu funkce $f(x)$ a jeden (libovolný) Clarkeův subgradient $g \in \partial f(x)$. Tato metoda pak na základě znalosti tohoto jednoho subgradientu vygeneruje (jednotkový) směr kroku d pro stávající iteraci. Hlavní nevýhodami této metody je neexistence implementovatelné ukončující podmínky a neexistence „dobré“ strategie pro určování délky kroku t^k (nelze použít line search, směr d nemusí být směr poklesu). Tím je způsoben malý řád

Obr. 15.5: Graf $f(x) = |x|$ Obr. 15.6: Graf $f(x) = (x^2 + \varepsilon)^{\frac{1}{2}}$

konvergence metody. Subgradientní metoda je tedy metoda největšího spádu, která nemá line search, ale volba délky kroku je provedena předem. Dá se ukázat, že délka kroku t^k musí splňovat následující omezení:

$$t^k > 0 \text{ a } \lim t^k = 0 \quad (15.2)$$

a

$$\sum_{k=0}^{\infty} t^k = +\infty. \quad (15.3)$$

Základní algoritmus subgradientní metody můžeme implementovat takto.

Algoritmus 15.1 Subgradientní metoda.

Je dána funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

Krok 0. {Inicializace.}

Zvolme $\varepsilon > 0$ (přesnost), $x^0 \in \mathbb{R}^n$, polož $k = 0$

Krok 1. {Cyklus subgradientní metody.}

while $\|x^k - x^{k-1}\| \geq \varepsilon$

Vyber libovolné $g^k \in \partial f(x^k)$

$$d^k = -\frac{g^k}{\|g^k\|}$$

t^k je určeno a-priori dle (15.2) a (15.3)

(příkladem takovéto posloupnosti je např. $t^k = \frac{1}{k}$)

$$x^{k+1} = x^k + t^k d^k$$

$$k = k + 1$$

end while

Krok 2. {Dosad (přibližné) řešení.}

$$\tilde{x} = x^k$$

15.4 Svazkové metody

Další širokou skupinou algoritmů, které jsou určeny pro minimalizaci nehladkých funkcí jsou bundle metody a jejich modifikace. Základní verze bundle metod byla vyvinuta nezávisle C. Lemarechalem a P. Wolfem a publikována byla v roce 1975. Tato metoda potřebuje opět proceduru, která vrací v každém bodě $x \in \mathbb{R}^n$ funkční hodnotu $f(x)$ a jeden (libovolný) Clarkeův subgradient $g \in \partial f(x)$.

Narozdíl od subgradientní metody ale bundle metoda uchovává subgradientní informaci nejen z aktuální iterace x_k , ale též subgradients g_i z minulých iterací y_i , množinu indexů subgradientů z minulých iterací uchovávaných ve svazku označíme J_k . Na základě těchto informací pak bundle metoda aproximuje Clarkeův zobecněný gradient v aktuální iteraci, tj. metoda má „úplnější“ informaci o chování funkce. Tyto uložené subgradients g_i modelují funkci f v aktuálním bodě x_k s různou přesností. Tato přesnost, která udává vzdálenost $g_i \in \partial f(y_i)$ od Clarkeova zobecněného gradientu $\partial f(x_k)$, je dána u konvexních funkcí jako

$$\alpha(x_k, y_i) = \alpha_i^k = f(x_k) - (f(y_i) + g_i^T(x_k - y_i)).$$

U nekonvexních funkcí popisuje tuto vzdálenost funkce

$$\beta(x_k, y_i) = \beta_i^k = \max\{\alpha_i^k, \|x_k - y_i\|^2\}.$$

Díky znalosti aproximace Clarkeova zobecněného gradientu je snadné implementovat ukončující podmínku (s využitím věty 15.9). Pro směr poklesu generovaný bundle metodou v bodě x_k platí $d_k = -z_k$, přičemž z_k je určeno konvexní kombinací subgradientů ze svazku. Navíc platí, že čím je větší chyba subgradientu β_i^k (příp. α_i^k), tím méně ovlivňuje daný subgradient g_i směr poklesu d_k .

Bundle metoda má také vypracovanou strategii volby délky kroku - u základní verze je to klasický line search a tím má tato metoda zaručen vyšší řád konvergence než subgradientní metoda.

Bundle metoda se také umí vyrovnat s případným „špatným“ směrem d_k . Pokud směr d_k nalezený bundle metodou v bodě x_k na základě aproximovaného Clarkeova zobecněného gradientu nevede k dostatečnému poklesu funkční hodnoty, provedeme tzv. *null step*, tj. obohatíme náš model přidáním dalšího subgradientu v blízkosti bodu x_k a tím „vylepšíme“ aproximaci Clarkeova zobecněného gradientu v bodě x_k a zůstaneme v bodě $x_{k+1} = x_k$.

V opačném případě (při dostatečném poklesu funkční hodnoty ve směru d_k) provedeme tzv. *serious step*, tj. přejdeme do nového bodu $x_{k+1} = x_k + td_k$, kde velikost kroku t je generována pomocí line search.

Protože velikost počítačové paměti je omezená a také kvůli zajištění efektivnosti výpočtu je nutné omezit počet subgradientů, kterými modelujeme Clarkeův zobecněný gradient. K tomu slouží tzv. *reset*, tj. při překročení únosného počtu subgradientů vybereme pouze některé subgradients z původního svazku a zbylé z něj vymažeme. Lemarechalův algoritmus pro minimalizaci nehladkých nekonvexních funkcí

realizovaný v jazyce Fortran vedl k úloze vyhledávání směru ve tvaru kvadratického programu

$$\begin{cases} \min \frac{1}{2} \left\| \sum_{i \in J_k} \lambda_i g_i \right\|^2, \\ \text{s omezením} \\ \lambda_i \geq 0, \quad i \in J_k, \\ \sum_{i \in J_k} \lambda_i = 1, \\ \sum_{i \in J_k} \lambda_i \beta_i^k \leq \varepsilon_k. \end{cases} \quad (15.4)$$

Směr $d_k = - \sum_{i \in J_k} \lambda_i^k g_i$, kde $\lambda_i^k, i \in J_k$ je řešením kvadratického programu (15.4).

Slabým místem takto navrženého algoritmu je nutnost volit dopředu koeficient ε_k .

Další modifikaci bundle metod zajistilo využití cutting plane method. Na základě cutting plane method modifikujeme bundle metodu tak, že odstraníme nepříjemnou volbu ε_k a priori, tj. volbu velikosti oblasti, na které se aproximuje Clarkeův zobecněný gradient $\partial f(x_k)$. Cutting plane method převádí problém minimalizace nehladké nekonvexní funkce $\min_{x \in \mathbb{R}^n} f(x)$ na iterační problém

$$\min_d \hat{f}_k(x_k + d) + \frac{1}{2t_k} \|d\|^2,$$

kde

$$\hat{f}_k(x_k + d) = \max_{i \in J_k} \{f(y_i) + g_i^T(x_k + d - y_i)\} = f(x_k) + \max_{i \in J_k} \{g_i^T d - \beta_i^k\}.$$

Řešení dané úlohy označíme d_k a toto d_k je opět směr poklesu, ve kterém provedeme line search následovaný serious stepem či null stepem stejně jako v předchozím případě. Z numerických důvodů je výhodné řešit duální formulaci výše uvedené úlohy, tím dostaneme opět úlohu kvadratického programování, pro kterou existují rychlé řešiče.

Duální formulace úlohy vypadá takto

$$\begin{cases} \min \frac{1}{2} \left\| \sum_{i \in J_k} \lambda_i g_i \right\|^2 + \frac{1}{t_k} \sum_{i \in J_k} \lambda_i \beta_i^k, \\ \text{s omezením} \\ \lambda_i \geq 0, \quad i \in J_k, \\ \sum_{i \in J_k} \lambda_i = 1. \end{cases} \quad (15.5)$$

Směr poklesu d_k dostaneme jako $d_k = -t_k \sum_{i \in J_k} \lambda_i^k g_i$, kde $\lambda_i^k, i \in J_k$ je řešením kvadratického programu (15.5). Očekávaný pokles ve směru d_k , je modelován funkcí

$$v(t_k) = -t_k \left\| \sum_{i \in J_k} \lambda_i g_i \right\|^2 - \sum_{i \in J_k} \lambda_i \beta_i^k.$$

Problematickým místem takto navrženého algoritmu je nyní nutnost volit dopředu koeficient t_k . Toto slabé místo je odstraněno použitím strategie pro určování délky kroku, kterou používá trust region metoda (podrobněji v kapitole 7). Připomeňme, že hlavní myšlenka trust region metody je minimalizace funkce f pouze na okolí bodu x_k , ve kterém očekáváme „dobrou“ shodu funkce f s jejím modelem. Toto okolí, přesněji velikost tohoto okolí se mění v průběhu výpočtu. To je implementováno následujícím způsobem.

Algoritmus 15.2 Trust region přístup.

- 1) Jestliže (skutečný pokles $\leq k_1 \cdot$ očekávaný pokles), pak zvětšete velikost okolí, na kterém minimalizujete funkci
- 2) Jestliže (skutečný pokles $\geq k_2 \cdot$ očekávaný pokles), pak zmenšete velikost okolí, na kterém minimalizujete funkci
- 3) Jinak neměňte velikost okolí, na kterém minimalizujete funkci

Mezi parametry k_1 a k_2 je následující vztah $0 < k_2 \leq k_1 < 1$. Touto strategií nahrazujeme line search v předchozí verzi bundle metody.

Spojením této strategie určování délky kroku a bundle metody vzniklé na základě cutting plane method dostáváme tzv. bundle trust metodu. Její stručný kompletní algoritmus uvádíme níže.

Algoritmus 15.3 Algoritmus bundle trust metody.

Krok 0. {Inicializace.}

Zvolte počáteční bod $x_1 \in \mathbb{R}^n$ a parametry $T > 0$, $0 < m_1 < m_2 < 1$, $0 < m_3 < 1$, $\varepsilon \geq 0$ a omezení velikosti svazku $j_{max} \geq 3$

Krok 1. Spočítejte $f(x_1)$, $g_1 \in \partial f(x_1)$ a položte $y_1 = x_1$, $J_1 = \{1\}$ a $k = 1$.

Krok 2. {Vnitřní iterace}

Spočítejte x_{k+1} a g_{k+1} pomocí vnitřní iterace uvedené níže nebo zjistěte, že x_k je „téměř stacionární“ (v tom případě skončete)

Krok 3. Jestliže $|J_k| = j_{max}$, pak jděte do Krok 4., jinak položte $J = J_k$ a pokračujte do Krok 5.

Krok 4. {Reset}

Vyberte $J \subset J_k$ při splnění $|J| \leq j_{max} - 2$ a $\max\{i \mid i \in J_k, \beta_{k,i} = 0\} \in J$. Dále zavedme další index \tilde{k} a definujme s použitím agregací $z_k = \sum_{i \in J_k} \lambda_i^k g_i$, $\sigma = \sum_{i \in J_k} \lambda_i \beta_i^k$: $g_{\tilde{k}} = z_k$,

$\beta_{\tilde{k}}^k = \sigma$, $J = J \cup \{\tilde{k}\}$.

Krok 5. {Update}

Jestliže výstupem vnitřní iterace byl Serious Step, potom položte: $\alpha_i^{k+1} = \alpha_i^k + f(x_{k+1}) - f(x_k) - g_i^T d_k$ pro $i \in J$, $\beta_i^{k+1} = \max\{\alpha_i^{k+1}, \|x_{k+1} - y_i\|^2\}$, $\beta_{k+1}^{k+1} = 0$.

Jestliže výstupem vnitřní iterace byl Null Step, potom položte: $\beta_i^{k+1} = \beta_i^k$ pro $i \in J$, $\beta_{k+1}^{k+1} = \beta(x_k, y_{k+1})$.

Položte $J_{k+1} = J \cup \{k+1\}$ a pokračujte do Krok 1.

Algoritmus 15.4 **Vnitřní iterace** $x_k \rightarrow x_{k+1}$.

- 1) Vyberte $t^1 = t_{k-1}$. Položte $l^1 = 0$, $u^1 = T$ a $j = 1$.
- 2) Spočítejte řešení $(v^j, d^j) = (v(t^j), d(t^j))$ úlohy (15.5). Jestliže $\frac{1}{t^j} \|d^j\| \leq \varepsilon$ a $-\frac{1}{t^j} \|d^j\|^2 - v^k \leq \varepsilon$, potom skončete: x_k je „téměř stacionární“. Jinak položte $y^j = x_k + d^j$ a spočítejte $g^j \in \partial f(y^j)$.
- 3)
 1. Jestliže je splněna podmínka $f(y^j) - f(x_k) < m_1 v^j$, potom proveďte Serious Step, tj. $x_{k+1} = y_{k+1} = y^j$, $g_{k+1} = g^j$ a skončete.
 2. Jestliže je splněna podmínka $f(y^j) - f(x_k) \geq m_1 v^j$ a současně $\alpha(x_k, y^j) \leq m_3 \sigma_{k-1}$ nebo $|f(x_k) - f(y^j)| \leq \|z_{k-1}\| + \sigma_{k-1}$ a ještě podmínka $(g^j)^T d^j - \beta_{k,j} \geq m_2 v^j$, potom proveďte Null Step, tj. $x_{k+1} = x_k$, $y_{k+1} = y^j$, $g_{k+1} = g^j$ a skončete.
 3. Jestliže je splněna podmínka $f(y^j) - f(x_k) \geq m_1 v^j$ a současně $\alpha(x_k, y^j) \leq m_3 \sigma_{k-1}$ nebo $|f(x_k) - f(y^j)| \leq \|z_{k-1}\| + \sigma_{k-1}$, ale není splněna podmínka $(g^j)^T d^j - \beta_{k,j} \geq m_2 v^j$, potom:
 - (a) Jestliže platí $|f(x_k) - f(y^j)| \leq \|z_{k-1}\| + \sigma_{k-1}$, potom položte $d_k = d^j$, $v_k = v^j$ a proveďte line search podél $x_k + s d_k$, $s \geq 0$.
 - (b) Jinak položte $u^{j+1} = t^j$, $l^{j+1} = l^j$, $t^{j+1} = \frac{1}{2}(u^{j+1} + l^{j+1})$, $j = j + 1$ a vraťte se do 2.
4. Jestliže je splněna podmínka $f(y^j) - f(x_k) \geq m_1 v^j$, ale není splněna podmínka $\alpha(x_k, y^j) \leq m_3 \sigma_{k-1}$ nebo $|f(x_k) - f(y^j)| \leq \|z_{k-1}\| + \sigma_{k-1}$ položte $u^{j+1} = t^j$, $l^{j+1} = l^j$, $t^{j+1} = \frac{1}{2}(u^{j+1} + l^{j+1})$, $j = j + 1$ a vraťte se do 2.

Nyní se ještě v krátkosti podívejme na konvergenci bundle trust metody pro nekonvexní a lokálně lipschitzovské funkce na \mathbb{R}^n v případě, že v předchozím algoritmu platí $\varepsilon = 0$.

Věta 15.15. *Nechť funkce f je slabě polohladká a lokálně lipschitzovsky spojitá na \mathbb{R}^n . Jestliže funkce f je omezená zdola a posloupnost iterací generována bundle trust metodou $\{x_k\}_{k \in \mathbb{N}}$ je omezená, pak existuje hromadný bod \bar{x} posloupnosti $\{x_k\}_{k \in \mathbb{N}}$ takový, že platí $0 \in \partial f(\bar{x})$.*

Kapitola 16

Úvod do metod globální optimalizace

V této kapitole se zaměříme z metod globální optimalizace na Nelder-Meadovu modifikaci základní simplexové metody. Tato metoda bývá obvykle označována jako Nelder-Meadova metoda. Jde o metodu, která obdobně jako jiné metody globální optimalizace, nepoužívají derivace (či gradienty) minimalizovaných funkcí ani aproximace těchto derivací (či gradientů). Tyto metody můžeme tedy obdobně jako v kapitole 3 označit jako metody bez derivací (derivative-free methods).

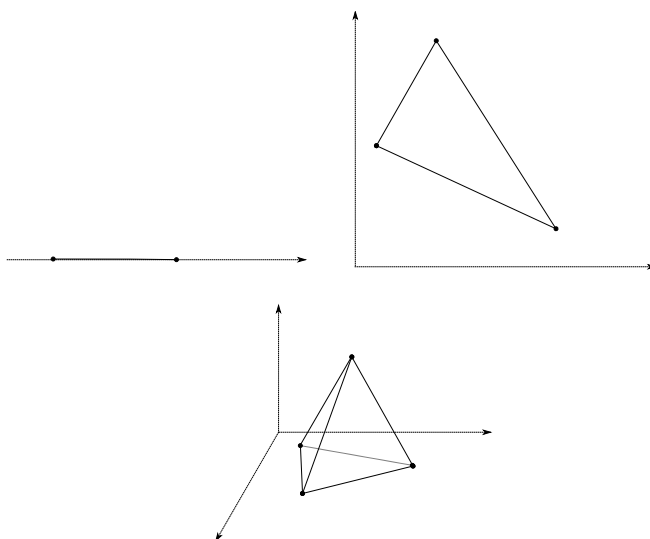
Hlavní výhodou těchto metod je, jak jsme již zmínili i v kapitole 3, je velmi snadná implementace a technická realizace těchto metod. Velká výhodou je v tom, že není třeba hledat ani Hessián ani gradient během iteračního procesu. Další výhodou je, že zatímco metody prvního a druhého řádu selhávají mimo jiné při minimalizaci nediferencovatelných či nespojitých funkcí, metody globální optimalizace jsou schopny najít minima i těchto funkcí. Z těchto důvodů jsou tyto metody stále v určité míře používány, i přes existenci daleko pokročilejších optimalizačních algoritmů.

16.1 Nelder-Meadova metoda simplexů

Nyní se budeme věnovat metodě vhodné pro vícerozměrnou optimalizaci. K tomu si potřebujeme zavést tzv. *simplex*. Simplex je zjednodušeně zobecněním trojúhelníku do libovolného n -rozměrného prostoru. Příklady simplexu v jedno, dvou a trojrozměrném prostoru jsou na obrázku 16.1.

Definice 16.1. Simplex či n -simplex je konvexní obal množiny $n + 1$ afinně nezávislých bodů v eukleidovském prostoru dimenze n nebo vyšší.

Nelder-Meadova metoda při minimalizaci v n -rozměrném prostoru pracuje se simplexem n bodů. V tomto simplexu se metoda snaží nahradit bod s „nejhorší“ (největší) funkční hodnotou novým bodem s „lepší“ (menší) funkční hodnotou. Tento

Obr. 16.1: Simplexy v \mathbb{R}^1 , \mathbb{R}^2 a \mathbb{R}^3

nový bod získáme buď pomocí tzv. zrcadlení, prodloužení či zkrácení simplexu podél přímky dané „nejhorším“ vrcholem a těžištěm zbylých bodů simplexu. Pokud tímto nenajdeme „lepší“ bod, pak zachováme bod s „nejlepší“ (nejmenší) funkční hodnotou a zúžíme celý simplex posunutím všech zbylých vrcholů simplexu směrem k tomuto bodu. Podrobně je uvedený postup ilustrován v následujícím algoritmu (algoritmus 16.1).

Algoritmus 16.1 Nelder-Meadova metoda.

```

1)  $\varepsilon > 0$  (přesnost),  $S := \text{simplex}$  ( $n+1$  bodů) z  $\mathbb{R}^n$ 
2) while  $f(z_{n+1}) - f(z_1) \geq \varepsilon$ 
    i) sestrojme  $z(-1)$  a spočtěme  $f_{-1} = f(z(-1))$  (zrcadlení)
    ii) if  $f(z_1) \leq f_{-1} \leq f(z_n)$ 
        (a)  $z_{n+1} := z(-1)$ 
    iii) elseif  $f_{-1} < f(z_1)$ 
        (a) sestrojme  $z(-2)$  a spočtěme  $f_{-2} = f(z(-2))$  (prodloužení)
        (b) if  $f_{-2} < f_{-1}$ 
            (c)  $z_{n+1} := z(-2)$ 
        (d) else
            (e)  $z_{n+1} := z(-1)$ 
        (f) end
    iv) elseif  $f_{-1} \geq f(z_n)$ 
        (a) if  $f(z_n) \leq f_{-1} \leq f(z_{n+1})$ 
            (b) sestrojme  $z(-\frac{1}{2})$  a spočtěme  $f_{-\frac{1}{2}} = f(z(-\frac{1}{2}))$  (vnější zkrácení)
            (c) if  $f_{-\frac{1}{2}} < f_{-1}$ 
                (d)  $z_{n+1} := z(-\frac{1}{2})$ 
            (e) end
            (f) else
                (g) sestrojme  $z(\frac{1}{2})$  a spočtěme  $f_{\frac{1}{2}} = f(z(\frac{1}{2}))$  (vnitřní zkrácení)
                (h) if  $f_{\frac{1}{2}} < f_{n+1}$ 
                    (i)  $z_{n+1} := z(\frac{1}{2})$ 
                (j) else
                    (k)  $z_i := \frac{1}{2}(z_1 + z_i)$ ,  $i = 2, 3, \dots, n+1$  (zúžení simplexu směrem k  $z_1$ )
                (l) end
            (m) end
        (v) end
3) end
4)  $z_1$  aproximuje minimum funkce  $f(x)$ 

```

z_{n+1} je bod, který má mezi body simplexu S největší hodnotu. z_1 je bod, který má mezi body simplexu S nejmenší hodnotu ($f(z_1) \leq f(z_2) \leq \dots \leq f(z_{n+1})$). \bar{z} je těžiště bodů z_1, z_2, \dots, z_n , $z(t) = \bar{z} + t(z_{n+1} - \bar{z})$ jsou body na přímce dané body

z_{n+1} a \bar{z} .

Předchozí Nelder-Meadova metoda se dá modifikovat pomocí řízeného náhodného výběru (Controlled Random Search). Parametr t , který určuje nový bod simplexu $z(t)$ je v tomto případě volen náhodně. Touto modifikací dostaneme následující algoritmus (Algoritmus 16.2). Mnohem více informací k tomuto algoritmu, který je použitelný pro minimalizaci funkcí s „menším“ počtem proměnných, u nichž je složité získat gradientní informaci, lze nalézt v [8].

Algoritmus 16.2 Nelder-Meadova metoda - úprava MCRS (Modified Controlled Random Search).

- 1) $\varepsilon > 0$ (přesnost), $\alpha \in \langle 4, 8 \rangle$, $P :=$ populace N náhodně vybraných bodů z \mathbb{R}^n
- 2) **while** $f(z_{max}) - f(z_{min}) \geq \varepsilon$
 - i) $S :=$ náhodný simplex ($n+1$ bodů) z populace P
 - ii) $z := g - t(x - g)$, kde x je náhodně zvolený bod ze simplexu S , g je těžiště zbylých bodů simplexu a $t \in \langle 0, \alpha \rangle$ je vybráno náhodně
 - iii) **if** $f(z) \leq f(z_{max})$
 - (a) $z_{max} := z$
 - iv) **end**
- 3) **end**
- 4) z_{min} aproximuje minimum funkce $f(x)$

z_{max} je bod, který má mezi N body populace P největší hodnotu. z_{min} je bod, který má mezi N body populace P nejmenší hodnotu.

Literatura

- [1] D. P. Bertsekas: *Nonlinear Programming* Athena Scientific (1999).
- [2] J. Frédéric Bonnans, Jean Charles Gilbert, Claude Lemaréchal, Claudia A. Sagastizábal: *Numerical Optimization: Theoretical and Practical Aspects* Springer (2006).
- [3] F. H. Clarke: *Optimization and Nonsmooth Analysis* Wiley (1983).
- [4] O. Došlý: *Základy konvexní analýzy a optimalizace v \mathbb{R}^n* Masarykova univerzita v Brně (2005).
- [5] Z. Dostál: *Optimal quadratic programming algorithms: with applications to variational inequalities* Springer (2009).
- [6] Z. Dostál, V. Vondrák: *Lineární algebra*. Text vytvořený při realizaci projektu „Matematika pro inženýry 21. století“, Vysoká škola báňská - Technická univerzita Ostrava (2012). (<http://mi21.vsb.cz/modul/linearni-algebra>)
- [7] R. Fletcher: *Practical Methods of Optimization* Wiley (1997).
- [8] J. Haslinger, D. Jedelský, T. Kozubek, J. Tvrdlík: *Genetic and Random Search Methods in Optimal Shape Design Problems* Journal of Global Optimizations 16 (2000).
- [9] T. Kozubek, T. Brzobohatý, V. Hapla, M. Jarošová, A. Markopoulos: *Lineární algebra s Matlabem*. Text vytvořený při realizaci projektu „Matematika pro inženýry 21. století“, Vysoká škola báňská - Technická univerzita Ostrava (2012). (<http://mi21.vsb.cz/modul/linearni-algebra-s-matlabem>)
- [10] J. Kuben, Š. Mayerová, P. Račková, P. Šarmonová: *Diferenciální počet funkcí více proměnných*. Text vytvořený při realizaci projektu „Matematika pro inženýry 21. století“, Vysoká škola báňská - Technická univerzita Ostrava (2012). (<http://mi21.vsb.cz/modul/diferencialni-pocet-funkci-vice-promennych>)
- [11] R. M. Lewis, V. Torczon, M. W. Trosset: *Direct search methods: then and now* Journal of Computational and Applied Mathematics 124 (2000).

-
- [12] J. Nocedal, S. Wright: *Numerical Optimization*. Springer (2006).
- [13] R. T. Rockafellar, R. J-B. Wets: *Variational Analysis*. Springer (1998).
- [14] P. Venkataraman: *Applied Optimization with MATLAB Programming*. Wiley (2009).
- [15] V. Vondrák, L. Pospíšil: *Numerické metody 1*. Text vytvořený při realizaci projektu „Matematika pro inženýry 21. století“, Vysoká škola báňská - Technická univerzita Ostrava (2012).
(<http://mi21.vsb.cz/modul/numerick%C3%A9-metody-1>)

Rejstřík

A

A -konjugované vektory, 32
 aktivní množina, 83, 85
 aktivní nerovnostní vazby, 83
 aproximace gradientu, 27, 28
 Armijovo pravidlo, 26, 43

B

bariérová funkce, 103
 bariérová metoda, 104, 105
 BFGS, 54, 59
 bisekce, 15–17
 bundle metoda, 133
 bundle trust metoda, 136, 137

C

cenová funkce, 1
 cenový funkcionál, 1
 centrální diference, 28
 Clarkeova zobecněná derivace, 126
 Clarkeův stacionární bod, 128
 Clarkeův subgradient, 127
 Clarkeův tečný kužel, 130
 Clarkeův zobecněný gradient, 127
 Constraint Qualifications, 100
 CQ, 100

D

dopředné diference, 28
 dualita
 pro kvadratické programování
 s nerovnostními vazbami, 110
 s rovnostními vazbami, 108
 pro rovnosti, 109
 pro nerovnosti, 112
 v konvexním programování, 108

duální funkce, 109
 duální problém
 pro rovnosti, 109
 pro nerovnosti, 112
 délka kroku, 26

F

Fibonacciova posloupnost, 18
 Fritz-Johnova podmínka, 91, 100
 funkce
 duální, 109
 konvexní, 11
 ryze konvexní, 11

G

globalizační strategie, 42
 globální optimalizace, 5
 globální řešení, 1
 gradientní metoda, 22–25
 Gramm–Schmidtův proces, 34

Ch

Choleského rozklad, 45, 56

I

isoperimetrický problém, 2

K

Karush-Kuhn-Tuckerova dvojice, 66, 68,
 77, 92, 100
 Karush-Kuhn-Tuckerova podmínka
 pro úlohy s omezením ve tvaru ne-
 rovností, 88, 92
 pro úlohy s omezením ve tvaru rov-
 ností, 66, 68, 72, 73
 pro úlohy s omezením ve tvaru rov-
 ností a nerovností, 100

- kladný poloprostor, 113
klasifikace omezení, 100
 na linearitu/konkávnost, 101
 na lineární nezávislost, 69, 91
konečné diference, 27, 28
konjugované vektory, 31, 32, 34
konvexní
 funkce, 11
 hranice, 11
 kombinace, 11
 množina, 10
 obal, 11
konvexní programování, 3
Krylovovův prostor, 34, 35
kužel, 67, 85
 polyhedrální, 85
kužel přípustných směrů, 89, 98
kužel směrů poklesu, 89
kvadratické programování, 4
kvazi-Newtonova metoda, 54
kvaziregulární bod, 90, 99
- L**
Lagrangeova funkce, 63, 64, 84
Lagrangian, 63, 64, 84
LICQ, 69, 91
Linear Independence Constraint Qualification, 69, 91
Linear/Concave Constraint Qualification, 101
linesearch, 14, 26
Lineární programování, 4
lineární programování, 115
lipschitzovská spojitost
 lokální, 126, 127
 na množině, 126
 v okolí bodu, 126, 127
lokální řešení, 3
- M**
Mangasarian-Fromovitz Constraint Qualification, 101
Mangasarian-Fromovitzova klasifikace omezení, 101
matice
 pásová, 46
MCRS metoda, 141
metoda
 trust region, 49, 135
 Levenberg-Marquardtova implementace, 51
metoda bariéry, 104, 105
metoda BFGS, 54, 59
metoda bundle, 133
metoda kvadratické penalty, 75, 105
metoda kvazi-Newtonova, 54
metoda multiplikátorů, 79
metoda největšího spádu, 22–25
metoda Nelder-Meadova, 139, 141, 142
metoda nepřímá, 131
metoda Newtonova, 42, 43
 s modifikací Hessiánu, 44
metoda nulového prostoru pro kvadratické programování, 78
metoda oboru hodnot pro kvadratické programování, 77
metoda rozšířených (augmented) Lagrangianů, 79
metoda sdružených gradientů, 31
 pro kvadratické funkce, 34, 37
 pro nelineární funkce, 39, 40
 Fletcher-Reeves, 39, 40
 Polak-Ribiere, 39, 40
metoda sdružených směrů
 pro kvadratické funkce, 31, 33
metoda sečen, 54
metoda simplexová, 118
metoda SMALBE, 106
metoda SMALE, 81
metoda subgradientní, 131, 132
metoda svazková, 133
metoda Uzawova typu, 78
metoda vnitřního bodu, 104, 117
metoda vnější (externí) penalty, 75, 105
metody globální optimalizace, 139
MFCQ, 101
min-max problém, 109, 111

minimalizace bez derivací, 14, 139
minimalizace bez omezení, 4
minimalizace s omezením, 62, 83

N

nadrovina, 113
nehladká optimalizace, 5, 125
Nelder-Meadova metoda, 139, 141, 142
nelineární programování, 5
neomezená optimalizaci, 3
nepřímá metoda, 131
Newtonova metoda, 42, 43
 s modifikací Hessiánu, 44
Newtonův krok, 42, 54
normálový kužel, 69, 90
null step, 133
návrhové proměnné, 3

O

omezená optimalizaci, 3
optimalizace bez omezení, 3
optimalizace s omezením, 3

P

podmínka komplementarity, 88
podmínky minima
 pro dvojnásobně diferencovatelnou funkci, 9
 pro kvadratickou funkci, 8
 pro kvadratické úlohy s omezením
 ve tvaru nerovností, 84, 87
 ve tvaru rovností, 63, 65
 pro nelineární úlohy s omezením
 ve tvaru nerovností, 88, 90, 91
 ve tvaru rovností, 66
 pro úlohy s omezením
 ve tvaru nerovností, 84
 ve tvaru rovností, 62, 63
polohladká funkce, 130
polyhedrální kužel, 85
polyhedrální množiny, 84
projekce na konvexní množinu, 85
přípustná křivka, 67
přípustná množina, 1, 3

přípustný směr, 66

R

regularita
 pro nerovnostní vazby, 91
 pro rovnostní vazby, 69
 pro rovnostní a nerovnostní vazby, 99
reset, 133
rozšířený Lagrangián, 75

S

SCQ, 102
serious step, 133
simplex, 139
simplexová metoda, 118
slabě polohladká funkce, 130
Slater Constraint Qualification, 102
Slaterova klasifikace omezení, 102
Snellův zákon lomu, 72, 74
spojitá optimalizace, 4
subdiferenciál, 127
subgradientní metoda, 131, 132
svazková metoda, 133

T

tečný kužel, 67, 88
Trust region, 49, 135
 Levenberg-Marquardtova implementace, 51

U

ukončovací podmínky, 28
unimodální funkce, 15
Uzawův algoritmus, 79

V

věta Farkasova, 85, 86
věta o lokálním extrému nehladké funkce, 127
věta o podmínce minima kvadratické funkce, 8
věta o podmínce minima pro dvojnásobně diferencovatelnou funkci, 9

věta Rademacherova, 126

věta Taylorova, 6, 7

W

Wolfeho podmínka, 57

Z

zlatý řez, 18–20

záporný poloprostor, 113