Aplikovaná Algebra (cvičení)

Oldřich Vlach, Zdeněk Dostál

20. dubna 2016

1 Připomenutí některých kapitol z lineární algebry

1.1 Vektorové prostory

Čtveřici $(\mathcal{V}, F, \oplus, \odot)$, kde $\mathcal{V}$ je množina, F je těleso [lépe řečeno $(F, +, \cdot)$ je těleso] a $\oplus$, $\odot$ jsou operace $\oplus : \mathcal{V} \times \mathcal{V} \mapsto \mathcal{V}$, $\odot : F \times \mathcal{V} \mapsto \mathcal{V}$ nazveme pojmem vektorový prostor, právě tehdy když: $(\mathcal{V}, \oplus)$ je komutativní grupa [vlastnosti (1)–(4)] a jsou splněny vlastnosti (5)–(8). Tedy přehledně

$$\begin{aligned}
\text{Asociativita: } (\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}) : &\quad (\mathbf{u} \oplus \mathbf{v}) \oplus \mathbf{w} = \mathbf{u} \oplus (\mathbf{v} \oplus \mathbf{w}) && (1)\\
\text{Nulový prvek: } (\exists \mathbf{o} \in \mathcal{V})(\forall \mathbf{u} \in \mathcal{V}) : &\quad \mathbf{u} \oplus \mathbf{o} = \mathbf{o} \oplus \mathbf{u} = \mathbf{u} && (2)\\
\text{Opačný prvek: } (\forall \mathbf{u} \in \mathcal{V})(\exists (-\mathbf{u}) \in \mathcal{V}) : &\quad \mathbf{u} \oplus (-\mathbf{u}) = \mathbf{o} && (3)\\
\text{Komutativita: } (\forall \mathbf{u}, \mathbf{v} \in \mathcal{V}) : &\quad \mathbf{u} \oplus \mathbf{v} = \mathbf{v} \oplus \mathbf{u} && (4)\\
\text{Distributivita k } \oplus : (\forall \alpha \in F)(\forall \mathbf{u}, \mathbf{v} \in \mathcal{V}) : &\quad \alpha \odot (\mathbf{u} \oplus \mathbf{v}) = (\alpha \odot \mathbf{u}) \oplus (\alpha \odot \mathbf{v}) && (5)\\
\text{Distributivita k } + : (\forall \alpha, \beta \in F)(\forall \mathbf{u} \in \mathcal{V}) : &\quad (\alpha + \beta) \odot \mathbf{u} = (\alpha \odot \mathbf{u}) \oplus (\beta \odot \mathbf{u}) && (6)\\
\text{Kompatibilita } \odot \text{ a } \cdot : (\forall \alpha, \beta \in F)(\forall \mathbf{u} \in \mathcal{V}) : &\quad \alpha \odot (\beta \odot \mathbf{u}) = (\alpha \cdot \beta) \odot \mathbf{u} && (7)\\
\text{Násobení jedničkou } F\text{: } (\forall \mathbf{u} \in \mathcal{V}) : &\quad 1 \odot \mathbf{u} = \mathbf{u} && (8)
\end{aligned}$$

kde $1 \in F$ je v tělese F identitou vzhledem k násobení, tedy $(\forall \alpha \in F)$: $1 \cdot \alpha = \alpha$.

Mějme podmnožinu $\mathcal{W} \subset \mathcal{V}$ a restrikce $\oplus_{|\mathcal{W}}$, $\odot_{|\mathcal{W}}$ operací $\oplus$, $\odot$ na $\mathcal{W}$. Pokud je $(\mathcal{W}, F, \oplus_{|\mathcal{W}}, \odot_{|\mathcal{W}})$ vektorovým prostorem, nazýváme ji pojmem podprostor vektorového prostotu $\mathcal{V}$. Postačující podmínkou, aby $\mathcal{W}$ byl podprostorem vektorového prostoru $\mathcal{V}$ je uzavřenost operací na $\mathcal{W}$

$$(\forall \mathbf{u}, \mathbf{v} \in \mathcal{W}) : \mathbf{u} \oplus_{|\mathcal{W}} \mathbf{v} \in \mathcal{W}$$
$$(\forall \alpha \in \mathbb{R})(\forall \mathbf{u}, \mathbf{v} \in \mathcal{W}) : \alpha \odot_{|\mathcal{W}} \mathbf{v} \in \mathcal{W}.$$

Příklad 1. Dokažte (4) a (5) pro

a) $(\mathbb{R}^{2\times 2}, \mathbb{R}, \oplus_{\mathbb{R}^{2\times 2}}, \odot_{\mathbb{R}^{2\times 2}})$

$$\begin{bmatrix} a & b \\ d & c \end{bmatrix} \oplus_{\mathbb{R}^{2\times 2}} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ d+g & c+h \end{bmatrix}$$
$$\alpha \odot_{\mathbb{R}^{2\times 2}} \begin{bmatrix} a & b \\ d & c \end{bmatrix} = \begin{bmatrix} \alpha a & \alpha b \\ \alpha d & \alpha c \end{bmatrix}$$

b) $(P_3, \mathbb{R}, \oplus_{P_3}, \odot_{P_3})$

$$(a + bx + cx^2) \oplus_{P_3} (d + ex + fx^2) = (a+d) + (b+c)x + (c+f)x^2$$
$$\alpha \odot_{P_3} (a + bx + cx^2) = (\alpha a + \alpha bx + \alpha cx^2)$$

Řesení.

a) $(\forall \alpha \in \mathbb{R})(\forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^{2\times 2})$:

$$\begin{aligned}
&\mathbf{u} = \begin{bmatrix} a & b \\ c & d \end{bmatrix},\ \mathbf{v} = \begin{bmatrix} e & f \\ g & h \end{bmatrix} \\
&\mathbf{u} \oplus_{\mathbb{R}^{2\times 2}} \mathbf{v} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \oplus_{\mathbb{R}^{2\times 2}} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix} = \\
&\quad = \begin{bmatrix} e+a & f+b \\ g+c & h+d \end{bmatrix} = \begin{bmatrix} e & f \\ g & h \end{bmatrix} \oplus_{\mathbb{R}^{2\times 2}} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \mathbf{v} \oplus_{\mathbb{R}^{2\times 2}} \mathbf{u} \\
&\alpha \odot_{\mathbb{R}^{2\times 2}} (\mathbf{u} \oplus_{\mathbb{R}^{2\times 2}} \mathbf{v}) = \alpha \odot_{\mathbb{R}^{2\times 2}} \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \oplus_{\mathbb{R}^{2\times 2}} \begin{bmatrix} e & f \\ g & h \end{bmatrix} \right) = \alpha \odot_{\mathbb{R}^{2\times 2}} \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix} = \\
&\quad = \begin{bmatrix} \alpha(a+e) & \alpha(b+f) \\ \alpha(c+g) & \alpha(d+h) \end{bmatrix} = \begin{bmatrix} \alpha a+\alpha e & \alpha b+\alpha f \\ \alpha c+\alpha g & \alpha d+\alpha h \end{bmatrix} = \begin{bmatrix} \alpha a & \alpha b \\ \alpha c & \alpha d \end{bmatrix} \oplus_{\mathbb{R}^{2\times 2}} \begin{bmatrix} \alpha e & \alpha f \\ \alpha g & \alpha h \end{bmatrix} = \\
&\quad = \left(\alpha \odot_{\mathbb{R}^{2\times 2}} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) \oplus_{\mathbb{R}^{2\times 2}} \left(\alpha \odot_{\mathbb{R}^{2\times 2}} \begin{bmatrix} e & f \\ g & h \end{bmatrix} \right) = (\alpha \odot_{\mathbb{R}^{2\times 2}} \mathbf{u}) \oplus_{\mathbb{R}^{2\times 2}} (\alpha \odot_{\mathbb{R}^{2\times 2}} \mathbf{v})
\end{aligned}$$

b) $(\forall \alpha \in \mathbb{R})(\forall \mathbf{u}, \mathbf{v} \in P_3)$:

$$\begin{aligned}
&\mathbf{u} = (a + bx + cx^2),\ \mathbf{v} = (d + ex + fx^2) \\
&\mathbf{u} \oplus_{P_3} \mathbf{v} = (a + bx + cx^2) \oplus_{P_3} (d + ex + fx^2) = (a+d) + (b+e)x + (c+f)x^2 = \\
&\quad (d+a) + (e+b)x + (f+c)x^2 = (d + ex + fx^2) \oplus_{P_3} (a + bx + cx^2) = \mathbf{v} \oplus_{P_3} \mathbf{u} \\
&\alpha \odot_{P_3} (\mathbf{u} \oplus_{P_3} \mathbf{v}) = \alpha \odot_{P_3} \big((a + bx + cx^2) \oplus_{P_3} (d + ex + fx^2) \big) = \\
&\quad = \alpha \odot_{P_3} \big((a+d) + (b+e)x + (c+f)x^2 \big) = \\
&\quad = \alpha(a+d) + \alpha(b+e)x + \alpha(c+f)x^2 = (\alpha a + \alpha d) + (\alpha b + \alpha e)x + (\alpha c + \alpha f)x^2 = \\
&\quad = \big(\alpha a + \alpha bx + \alpha cx^2\big) + \big(\alpha d + \alpha ex + \alpha fx^2\big) = \\
&\quad = \big(\alpha \odot_{P_3} (a + bx + cx^2)\big) \oplus_{P_3} \big(\alpha \odot_{P_3} (d + ex + fx^2)\big) = (\alpha \odot_{P_3} \mathbf{u}) \oplus_{P_3} (\alpha \odot_{P_3} \mathbf{v})
\end{aligned}$$

$\square$

Příklad 2. Mějme vektorový prostor $(\mathbb{R}\to\mathbb{R},\mathbb{R},\oplus_{\mathbb{R}\to\mathbb{R}},\odot_{\mathbb{R}\to\mathbb{R}})$ všech reálných funkcí, kde

$$\begin{aligned} \mathbb{R}\to\mathbb{R} &= \{f:\mathbb{R}\to\mathbb{R}\} \\ \big(\forall f,g\in\mathbb{R}\to\mathbb{R}\big)\big(\forall x\in\mathbb{R}\big):\quad (f\oplus_{\mathbb{R}\to\mathbb{R}} g)(x) &= f(x)+g(x) \\ \big(\forall\alpha\in\mathbb{R}\big)\big(\forall f\in\mathbb{R}\to\mathbb{R}\big)\big(\forall x\in\mathbb{R}\big):\quad (\alpha\odot_{\mathbb{R}\to\mathbb{R}} f)(x) &= \alpha f(x). \end{aligned}$$

Určete (a své tvrzení obhajte), které z množin $\# \in \{\mathcal{U},\mathcal{V},\mathcal{W}\}\subset\mathbb{R}\to\mathbb{R}$

$$\begin{aligned} \mathcal{U} &= C^1(\mathbb{R},\mathbb{R}) = \{f:\mathbb{R}\to\mathbb{R}|\ (\forall x\in\mathbb{R}):\ f'(x)<\infty\} \\ \mathcal{U} &= \{f:\mathbb{R}\to\mathbb{R}|\ (\forall x\in\mathbb{R}):\ f(2016)=1\} \\ \mathcal{W} &= \{f:\mathbb{R}\to\mathbb{R}|\ (\forall x\in\mathbb{R}):\ f(x)=c\in\mathbb{Z}\} \end{aligned}$$

spolu s restrikcí operací z $C^1(\mathbb{R},\mathbb{R})$ na # tvoří podprostor vektorového prostoru $C^1(\mathbb{R},\mathbb{R})$.

Řesení. Je potřeba prokázat uzavřenost operací sčítání vektorů a násobení skalárem.

a) $\big(\mathcal{U}=C^1(\mathbb{R},\mathbb{R}),\mathbb{R},\oplus_{|\mathcal{U}},\odot_{|\mathcal{U}}\big)$ je podprostor vektorového prostoru $(\mathbb{R}\to\mathbb{R},\mathbb{R},\oplus_{\mathbb{R}\to\mathbb{R}},\odot_{\mathbb{R}\to\mathbb{R}})$ jelikož

$$\begin{aligned} f,g\in\mathcal{U}\ &\Rightarrow (\forall x\in\mathbb{R}):\ \big(f\oplus_{|\mathcal{U}} g\big)'(x)=f'(x)+g'(x)<\infty\ \Rightarrow\ f\oplus_{|\mathcal{U}} g\in\mathcal{U} \\ \alpha\in\mathbb{R},\ f\in\mathcal{U}\ &\Rightarrow (\forall x\in\mathbb{R}):\ \big(\alpha\odot_{|\mathcal{U}} f\big)'(x)=\alpha f'(x)<\infty\ \Rightarrow\ \alpha\odot_{|\mathcal{U}} f\in\mathcal{U}. \end{aligned}$$

b) $\big(\mathcal{V},\mathbb{R},\oplus_{|\mathcal{V}},\odot_{|\mathcal{V}}\big)$ není vektorový prostor, jelikož například dokonce $(\forall f,g\in\mathcal{V})$:

$$\big(f\oplus_{|\mathcal{V}} g\big)(2016)=f(2016)+g(2016)=1+1=2\ \Rightarrow\ f\oplus_{|\mathcal{V}} g\notin\mathcal{V}.$$

Podobně není uzavřené například násobení skalárem $\alpha\neq 1$. Všimněme si, že nulová funkce $\mathbf{0}_{\mathbb{R}\to\mathbb{R}}\equiv 0$ z axiomu 2 vektorového prostoru $\mathbb{R}\to\mathbb{R}$ není prvkem $\mathcal{V}$. Podprostor vektorového prostoru vždy musí obsahovat nulový prvek původního prostoru.

c) $\big(\mathcal{W},\mathbb{R},\oplus_{|\mathcal{W}},\odot_{|\mathcal{W}}\big)$ není vektorový prostor, jelikož pro například pro $\alpha=\frac{1}{2}$, $f\equiv 1$ je

$$(\forall x\in\mathbb{R}):\ \big(\alpha\odot_{|\mathcal{W}} f\big)(x)=\frac{1}{2}\cdot 1=\frac{1}{2}\notin\mathbb{Z}\ \Rightarrow\ \alpha\odot_{|\mathcal{W}} f\equiv\frac{1}{2}\notin\mathcal{W}.$$

Všimněme si, že operace $\oplus_{|\mathcal{W}}$ uzavřená je

$$(\forall f,g\in\mathcal{W}):\ (\forall x\in\mathbb{R}):\ \big(f\oplus_{|\mathcal{W}} g\big)(x)=f(x)+g(x)=c_f+c_g=c_{f\oplus_{|\mathcal{W}}g}\in\mathbb{Z},\ \text{tedy}\ f\oplus_{|\mathcal{W}} g\in\mathcal{W}.$$

$\square$

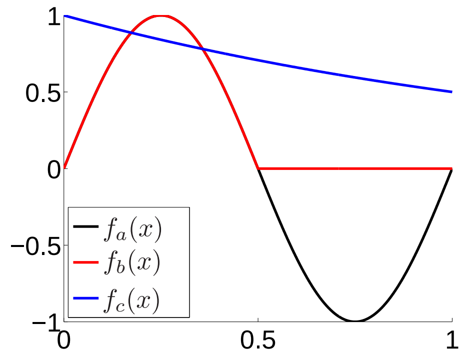
1.2 Báze

Uspořádanou množinu $\mathcal{F} = (\mathbf{f}_1, \dots, \mathbf{f}_n)$, vektorů vektorového prostoru $\mathcal{V}$ nezveme pojmem báze vektorového prostoru $\mathcal{V}$, jestliže je lineárně nezávislá a její lineární obal je celý prostor $\mathcal{V}$, tedy $\overline{\langle \mathcal{F} \rangle} = \mathcal{V}$. Souřadnice $[\mathbf{u}]_{\mathcal{F}}$ vektoru $\mathbf{u} \in \mathcal{V}$ vzhledem k bázi $\mathcal{F}$ vektorového prostoru $\mathcal{V}$ je vektor koeficientů lineární kombinace $\mathbb{R}^n \ni [\mathbf{u}]_{\mathcal{F}} :\ \mathbf{u} = ([\mathbf{u}]_{\mathcal{F}})_1\, \mathbf{f}_1 + \dots + ([\mathbf{u}]_{\mathcal{F}})_n\, \mathbf{f}_n$.

Příklad 3. Mějme několik reálných spojitých funkcí $f_\# \in C(\langle 0,1 \rangle, \mathbb{R})$, $\# \in \{a, b, c\}$, zadaných předpisy

$$f_a(x) = \sin(2\pi x), \quad f_b(x) = \begin{cases} \sin(2\pi x) & x \in \langle 0, \frac{1}{2}) \\ 0 & x \in (\frac{1}{2}, 1\rangle \end{cases}, \quad f_c(x) = 2^{-x}$$

a zobrazených na Obrázku 1. Chceme tyto funkce reprezentovat na počítači, což nelze udělat pro obecné



Obrázek 1: Funkce z Příkladu 3.

funkce přesně, jelikož prostor $C(\langle 0,1 \rangle, \mathbb{R})$ má nekonečnou dimenzi. Zkusme to nepřesně. Interval $(0, 1)$ rozdělíme na 2^n dílků (proč, to uvidíme později) a dostaneme rovnoměrně navzorkování intervalu $(0, 1)$ body x_j, $j \in 1, \dots, 2^n + 1$,

$$x_1 = 0,\ x_2 = 2^{-n}, \dots,\ x_j = (j-1) \cdot 2^{-n}, \dots,\ x_{2^n+1} = 1.$$

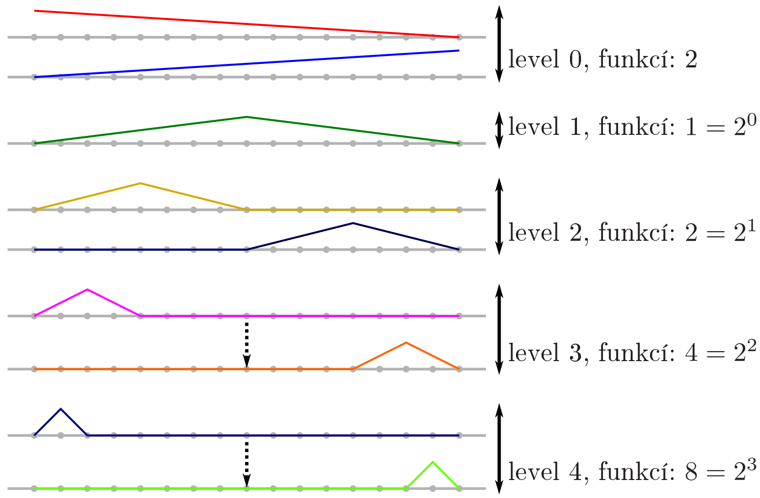
Hodnoty zadaných funkcí v těchto bodech označíme $f_{\#,j}$. Funkce $f_\#$ nahradíme funkcemi $\tilde{f}_\# \in \mathcal{V} \subset C(\langle 0,1 \rangle, \mathbb{R})$, kde funkce z prostoru $\mathcal{V}$ jsou po částech lineární na intervalech

$$I_j = \langle x_j, x_{j+1} \rangle, \quad j \in \{1, \dots, 2^n\} \quad \text{a} \quad f_\#(x_k) = \tilde{f}_\#(x_k), \quad k \in \{1, \dots, 2^n + 1\},\ \# \in \{a, b, c\}.$$

a) Určete v jaké bázi $\mathcal{S}$ prostoru $\mathcal{V}$ jsou platí pro souřadnice

$$[\tilde{f}_\#]_{\mathcal{S}} = \begin{bmatrix} \tilde{f}_\#(x_1) \\ \vdots \\ \tilde{f}_\#(x_{2^n+1}) \end{bmatrix}.$$

b) Rekurzivně definujme jinou bázi $\mathcal{E}$, viz obrázek 2. Určete souřadnice funkcí $\tilde{f}_\#$ v této nové bázi a proveďte tzv prahování, kdy všechny souřadnice menší než $\varepsilon = 10^{-2}$ zanedbejte.



Obrázek 2: Hierarchická báze na intervalu $\langle 0, 1\rangle$, první čtyři úrovně.

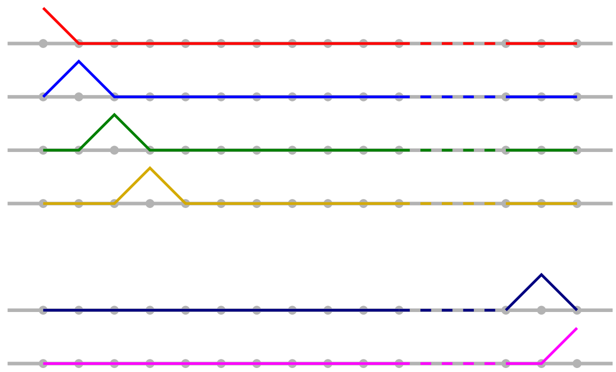
Listing 1: generate_1d_hierarchical_basis.m

```
function B = generate_1d_hierarchical_basis(n)
% generates the hierarchical basis on interval (0,1) of order n
% the basis is returned in columns of B that contains
% "dofs" functions b_{1} ... b_{dofs}
% each function is given by values in points x
dofs = 1+2^n;             % B contains "dofs" functions
x = (0:dofs-1)/(dofs-1); %
B = zeros(dofs);
B(:,1) = x;
B(:,2) = 1-x;
count = 3;
for i = 1:n
  % for remaining levels
  xa = 0         :2^(n-i+1):2^n-2^(n-i+1);
  xb = 2^(n-i+1):2^(n-i+1):2^n;
  xc = 0.5*(xa+xb);
  for j = 1:size(xa,2);
    B(xa(j)+1:xc(j)+1,count) =   (0:xc(j)-xa(j))'/(xc(j)-xa(j));
    B(xc(j)+1:xb(j)+1,count) = 1-(0:xb(j)-xc(j))'/(xb(j)-xc(j));
    count = count+1;
  end
end
```

Řesení. Zdrojové kódy k tomuto příkladu jsou uvedeny v 1 a 2.

a) Báze $\mathcal{S}$ je zobrazena na obrázku 3. Jedná se o po částech lineární spojité funkce φ_j, které mají ve vzorkovacím bodě x_j hodnotu 1, a v ostatních vzorkovacích bodech hodnotu 0. Je zřejmé, že souřadnice v této bázi jsou přímo funkční hodnoty $f_{\#}(x_k)$ a lze je tak ztotožnit s grafy funkcí na Obrázku 1.

b) Řešení tohoto úkolu je uvedeno ve zdrojovém kódu 2. Nejdříve se spustí generování hodnot bázových funkcí $\psi_j(x_i)$ ve vzorkovacích bodech a uloží se do matice $\mathsf{B} \in \mathbb{R}^{(2^n+1)\times(2^n+1)}$, tedy



Obrázek 3: Standardní báze na intervalu $\langle 0, 1 \rangle$.

$[\mathsf{B}]_{i,j} = \psi_j(x_i)$. Funkční hodnoty zadaných funkcí ve vzorkovacích bodech uložíme do matice $\mathsf{Y} \in \mathbb{R}^{(2^n+1)\times 3}$, podrobněji $[\mathsf{C}]_{i,j} = \tilde{f}_j(x_i)$.
Potom spočítáme souřadnice v bázi $\mathcal{E} = (\psi_1, \dots, \psi_{2^n+1})$, které uložíme po sloupcích do matice $\mathsf{C} \in \mathbb{R}^{(2^n+1)\times(2^n+1)}$

$$[\tilde{f}_a]_{\mathcal{E}} = [\mathsf{C}]_{:,1}, \; [f_b]_{\mathcal{E}} = [\mathsf{C}]_{:,2}, \; [f_c]_{\mathcal{E}} = [\mathsf{C}]_{:,3}.$$

Například pro první sloupec matice C pak musí platit

$$\tilde{f}_a = \sum_{i=1}^{2^n+1} [\mathsf{C}]_{i,1}\psi_i \;\Rightarrow\; [\mathsf{Y}]_{:,1} = \sum_{i=1}^{2^n+1} [\mathsf{C}]_{i,1}[\mathsf{B}]_{:,i} = \sum_{k=1}^{2^n+1} [\mathsf{B}]_{:,k}[\mathsf{C}]_{k,1} = \mathsf{B}\cdot[\mathsf{C}]_{:,1},$$

a analogicky to bude pro ostatní sloupce matice C. Tedy $\mathsf{Y} = \mathsf{BC}$, $\mathsf{C} = \mathsf{B}^{-1}\mathsf{Y}$. Podotkněme, že pro výpočet matice C není nutné (a ani vhodné, proč?) spočítat matici B^{-1}. Absolutní hodnoty získaných souřadnic jsou zobrazeny na Obrázku 4. Všimneme si, že podstatná většina z nich je velmi malá. Co se stane, když všechny menší než ε znulujeme? Zkopírujeme matici C do matice D a následně v D provedeme prahování hodnot na ε. Tímto procesem

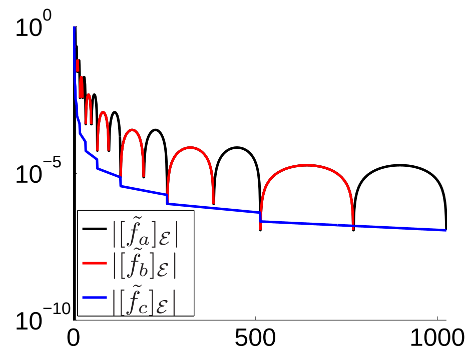
$$\hat{f}_a = \sum_{i=1}^{2^n+1} [\mathsf{D}]_{i,1}\psi_i$$

(a analogicky) obdržíme funkce $\hat{f}_\#$, které jsou zobrazeny na Obrázku 5 (vlevo) a rozdíly (chyba, kterou se dopouštíme prahováním souřadnic) $f_\# - \hat{f}_\#$ je zobrazena na Obrázku 5 (vpravo).
Ještě uvedeme, jak se prahování projeví na počtu nenulových koeficientů, což je uvedeno v Tabulce 1.

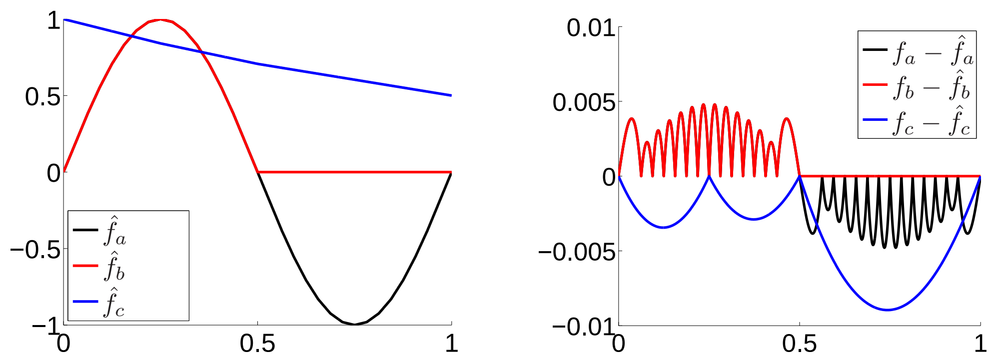
#	a	b	c
$f_\#$	1024	511	1025
$\hat{f}_\#$	26	13	4

Tabulka 1: Počty nenulových koeficientů po prahování, Příklad 3.

□



Obrázek 4: Absolutní hodnoty souřadnic v hierarchické bázi, Příklad 3.



Obrázek 5: Fuknce po prahování souřadnic, Příklad 3.

Listing 2: hierarchical_basis_test.m

```
n = 10;                                  % level
B = generate_1d_hierarchical_basis(n);   % hierarchical basis
m = size(B,1); z=1:m;                    % dofs
x = (0:m-1)/(m-1);                       % interval (0,1)
e = 1e-2;                                % epsilon
f = @(x) [...                            % test functions
  sin(2*pi*x);...
  sin(2*pi*x).*(x<0.5);...
  2.^(-x)
  ];
Y = f(x)';                       figure(1); hold on;
plot(x,    Y(:,1) ,'-k'); plot(x,    Y(:,2) ,'-r'); plot(x,    Y(:,3) ,'-b');
C = B\Y;                       figure(2); hold on;
plot(x,  B*C(:,1) ,'-k'); plot(x,  B*C(:,2) ,'-r'); plot(x,  B*C(:,3) ,'-b');
figure(3); hold on;
plot(z,    C(:,1) ,'-k'); plot(z,    C(:,2) ,'-r'); plot(z,    C(:,3) ,'-b');
figure(4); hold on;
plot(z,abs(C(:,1)),'-k'); plot(z,abs(C(:,2)),'-r'); plot(z,abs(C(:,3)),'-b');
set(gca,'YScale','log');
D = C; D(abs(D)<e) = 0;   figure(5); hold on;
plot(x,  B*D(:,1) ,'-k'); plot(x,  B*D(:,2) ,'-r'); plot(x,  B*D(:,3) ,'-b');
fprintf(1,['Nonzero valued coordinates (hierarchical); ',...
  'f1: %4i,  f2: %4i,  f3: %4i\n'],nnz(C(:,1)), nnz(C(:,2)), nnz(C(:,3)));
fprintf(1,['Nonzero valued coordinates (treshold h.);  ',...
  'f1: %4i,  f2: %4i,  f3: %4i\n'],nnz(D(:,1)), nnz(D(:,2)), nnz(D(:,3)));
```

1.3 Lineární zobrazení

Zobrazení $A: \mathcal{U} \to \mathcal{V}$ nazveme lineární (zapisujeme $A \in \mathscr{L}(\mathcal{U}, \mathcal{V})$) právě tehdy, když

$$
\begin{aligned}
&(\forall \mathbf{u}, \mathbf{v} \in \mathcal{U}): && A(\mathbf{u}) \oplus_{\mathcal{V}} A(\mathbf{v}) = A(\mathbf{u} \oplus_{\mathcal{U}} \mathbf{v}) && (9)\\
&(\forall \alpha \in \mathbb{R})(\forall \mathbf{u} \in \mathcal{U}): && \alpha \odot_{\mathcal{V}} A(\mathbf{u}) = A(\alpha \odot_{\mathcal{U}} \mathbf{u}). && (10)
\end{aligned}
$$

Nechť vektorové prostory $\mathcal{U}$ a $\mathcal{V}$ mají báze $\mathcal{E} = (\mathbf{e}_1, \ldots, \mathbf{e}_m)$ a $\mathcal{F} = (\mathbf{f}_1, \ldots, \mathbf{f}_n)$. Matice lineárního zobrazení $A \in \mathscr{L}(\mathcal{U}, \mathcal{V})$ vzhledem k bázím $\mathcal{E}$ a $\mathcal{F}$ je matice

$$
\mathbb{R}^{n \times m} \ni \mathsf{A}_{\mathcal{E},\mathcal{F}} \overset{\text{ozn.}}{\equiv} \mathsf{A}_{\mathcal{F} \leftarrow \mathcal{E}} = \left[[A(\mathbf{e}_1)]_{\mathcal{F}} \quad \cdots \quad [A(\mathbf{e}_1)]_{\mathcal{F}} \right].
$$

Platí pro ni $[A(\mathbf{u})]_{\mathcal{F}} = \mathsf{A}_{\mathcal{F} \leftarrow \mathcal{E}} [\mathbf{u}]_{\mathcal{E}}$. Připomínáme, že v tomto textu budeme označovat matici přechodu $\mathsf{A}_{\mathcal{F} \leftarrow \mathcal{E}}$ místo obvyklého $\mathsf{A}_{\mathcal{E},\mathcal{F}}$.

Příklad 4. Mějme n bodů v $\mathbb{R}^3$ (například souřadnice bodů nějakého STL modelu). Jak se změní tyto souřadnice při rotaci postupně o úhly $(\alpha, \beta, \gamma) = (10°, 15°, 20°)$ vzhledem k osám (x, y, z)?

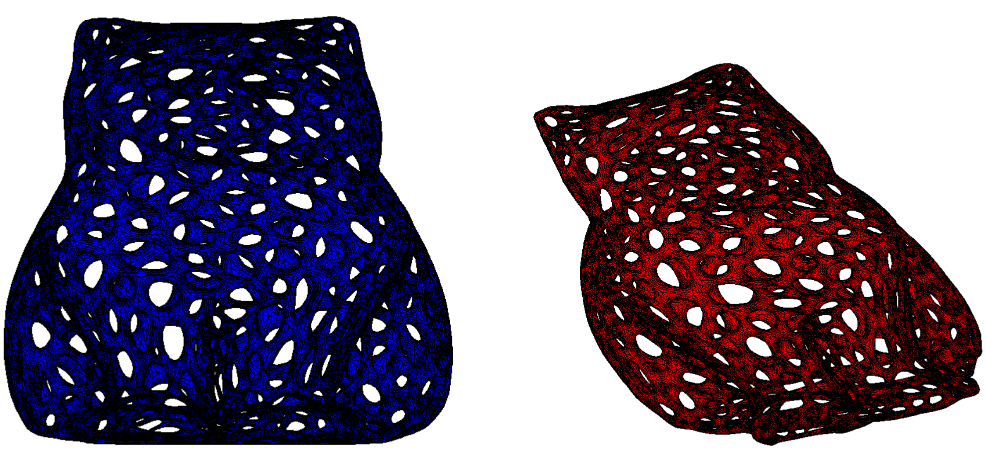
Řesení. Výslednou rotaci $R_{\alpha,\beta,\gamma} \in \mathscr{L}(\mathbb{R}^3, \mathbb{R}^3)$ dostaneme postupnou rotací (zobrazení $R_{x,\alpha}, R_{y,\beta}, R_{z,\gamma} \in \mathscr{L}(\mathbb{R}^3, \mathbb{R}^3)$) podél jednotlivých os, tedy $R_{\alpha,\beta,\gamma} = R_{x,\alpha} \circ R_{y,\beta} \circ R_{z,\gamma}$, kde

$$
[R_{x,\alpha}]_{\mathcal{S},\mathcal{S}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{bmatrix}, \quad [R_{y,\beta}]_{\mathcal{S},\mathcal{S}} = \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix}, \quad [R_{z,\gamma}]_{\mathcal{S},\mathcal{S}} = \begin{bmatrix} \cos(\beta) & -\sin(\beta) & 0 \\ \sin(\beta) & \cos(\beta) & 0 \\ 0 & 0 & 1 \end{bmatrix}.
$$

Zdrojové kódy k tomuto příkladu jsou uvedeny v Kódech **??** a **??**. Jako příklad použití si načteme ASCII STL soubor, obsahující povrchovou síť modelu kočky, funkcí `read_stl`, která vrací body P (matice $n_p \times 3$). Povrchovou síť zobrazíme pomocí funkce `patch`, viz Obrázek 6 (vlevo). Rotované souřadnice P_r získáme pomocí matice R, což je vlastně matice lineární transformace $R_{\alpha,\beta,\gamma} \in \mathscr{L}(\mathbb{R}^3, \mathbb{R}^3)$ vzhledem ke standardní bázi $\mathcal{S}$, to jest $\mathsf{R} = [R_{\alpha,\beta,\gamma}]_{\mathcal{S},\mathcal{S}}$. Proto

$$
\mathsf{P}_r^\top = \mathsf{R}\mathsf{P}^\top \;\Rightarrow\; \mathsf{P}_r = \mathsf{P}\mathsf{R}^\top.
$$

Na Obrázku 6 vpravo je zobrazena povrchová síť po rotaci.



Obrázek 6: Data z STL souboru vlevo originální, vpravo po rotaci, Příklad 4.

$\square$

Listing 3: read_stl.m

```
function [p,e,n] = read_stl(filename)
% reads STL ascii file
% outs: p ... points
%       e ... triangles
%       n ... outer normals to triangles
fid = fopen(filename,'r');
if fid
  % read stl file
  textscan(fid,'%s',1,'delimiter','\n');
  ts=textscan(fid,['facet normal %f %f %f\n','outer loop\n',...
    'vertex%f %f %f\n','vertex%f %f %f\n','vertex%f %f %f\n',...
    'endloop\n','endfacet']);
  n = [ts{ 1} ts{ 2} ts{ 3}];
  p = [ts{ 4} ts{ 5} ts{ 6}; ts{ 7} ts{ 8} ts{ 9}; ts{10} ts{11} ts{12}];
  m = size(n, 1);
  e = [1:m; m+1:2*m; 2*m+1:3*m];
end
```

Listing 4: rotation_test.m

```
[p,e,n] = read_stl('../data_files/Mesh.stl');
al = 10*pi/180;
be = 15*pi/180;
ga = 20*pi/180;
R2d = @(xx) [cos(xx) -sin(xx); sin(xx) cos(xx)];
Rx = eye(3); Rx([2 3],[2 3]) = R2d(al);
Ry = eye(3); Ry([1 3],[1 3]) = R2d(be)';
Rz = eye(3); Rz([1 2],[1 2]) = R2d(ga);
Rt = eye(3); Rt([2 3],[2 3]) = R2d(90*pi/180);
R = Rx*Ry*Rz; p = p*Rt;
pr = p*R';
figure(1); patch('Faces',e','Vertices',p ,'FaceColor','b');
figure(2); patch('Faces',e','Vertices',pr,'FaceColor','r');
```

Důležitou lineární transformací je změna souřadnic při změně báze. Mějme dvě báze $\mathcal{E}$, $\mathcal{F}$ vektorového prostoru $\mathcal{V}$. Pro každý vektor $\mathbf{v} \in \mathcal{V}$ umíme určit souřadnice $[\mathbf{v}]_{\mathcal{E}}$, $[\mathbf{v}]_{\mathcal{F}}$ v těchto bázích. Ukažme, že zobrazení $P_{\mathcal{F}\leftarrow\mathcal{E}}$ přiřazující souřadnicím v bázi $\mathcal{E}$ příslušné souřadnice v bázi $\mathcal{F}$ je lineární, tj. $P_{\mathcal{F}\leftarrow\mathcal{E}} \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^n)$, $n = \dim \mathcal{V}$, navíc umíme spočíst matici tohoto lineárního zobrazení $\mathsf{P}_{\mathcal{F}\leftarrow\mathcal{E}} \in \mathbb{R}^{n\times n}$, pomocí které pak snadno převádíme souřadnice

$$[\mathbf{v}]_{\mathcal{F}} = \mathsf{P}_{\mathcal{F}\leftarrow\mathcal{E}}[\mathbf{v}]_{\mathcal{E}}, \quad [\mathbf{v}]_{\mathcal{E}} = \mathsf{P}_{\mathcal{E}\leftarrow\mathcal{F}}[\mathbf{v}]_{\mathcal{F}} \;\Rightarrow\; \mathsf{P}_{\mathcal{F}\leftarrow\mathcal{E}}^{-1} = \mathsf{P}_{\mathcal{E}\leftarrow\mathcal{F}}\,.$$

Skutečně pro souřadnice platí

$$\mathbf{v} = \sum_{i=1}^{n} [\mathbf{v}]_{\mathcal{F},i}\mathbf{f}_i = \sum_{i=1}^{n} [\mathbf{v}]_{\mathcal{E},i}\mathbf{e}_i\,.$$

Tuto rovnost mezi vektory ve $\mathcal{V}$ vyjádříme v souřadnicích některé z bází, například $\mathcal{F}$ (vekory $\mathbf{s}_i$ jsou vekory standardní báze prostoru $\mathbb{R}^n$) a

$$[\mathbf{v}]_{\mathcal{F}} = \sum_{i=1}^{n} [\mathbf{v}]_{\mathcal{F},i} \overbrace{[\mathbf{f}_i]_{\mathcal{F}}}^{\mathbf{s}_i} = \sum_{i=1}^{n} [\mathbf{v}]_{\mathcal{E},i}[\mathbf{e}_i]_{\mathcal{F}} = \begin{bmatrix} [\mathbf{e}_1]_{\mathcal{F}} & \dots & [\mathbf{e}_n]_{\mathcal{F}} \end{bmatrix} \begin{bmatrix} [\mathbf{v}]_{\mathcal{E},1} \\ \vdots \\ [\mathbf{v}]_{\mathcal{E},n} \end{bmatrix} = \mathsf{P}_{\mathcal{F}\leftarrow\mathcal{E}}[\mathbf{v}]_{\mathcal{E}}\,.$$

Dostali jsme vztah pro tzv. <u>matici přechodu</u> od souřadnic v bázi $\mathcal{E}$ k souřadnicím v bázi $\mathcal{F}$

$$\mathsf{P}_{\mathcal{F}\leftarrow\mathcal{E}} = \begin{bmatrix} [\mathbf{e}_1]_{\mathcal{F}} & \dots & [\mathbf{e}_n]_{\mathcal{F}} \end{bmatrix}\,.$$

Příklad 5. Mějme vektorový prostor $(\mathbb{R}^{2\times 2}, \mathbb{R}, \oplus_{\mathbb{R}^{2\times 2}}, \odot_{\mathbb{R}^{2\times 2}})$, báze $\mathcal{E}$, $\mathcal{F}$, dále lineární zobrazení $A : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$ a bilineární formu $B : \mathbb{R}^{2\times 2} \times \mathbb{R}^{2\times 2} \to \mathbb{R}$

$$\mathcal{E} = \left(\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right)$$

$$\mathcal{F} = \left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} \right)$$

$$A(\mathbf{u}) = \begin{bmatrix} 8 & 6 \\ 4 & 2 \end{bmatrix} \mathbf{u}$$

$$B(\mathbf{u}, \mathbf{v}) = \begin{bmatrix} [\mathbf{u}]_{1,1} & [\mathbf{u}]_{1,2} & [\mathbf{u}]_{1,2} & [\mathbf{u}]_{2,2} \end{bmatrix} \begin{bmatrix} 11 & 12 & 13 & 14 \\ 21 & 22 & 23 & 24 \\ 31 & 32 & 33 & 34 \\ 41 & 42 & 43 & 44 \end{bmatrix} \begin{bmatrix} [\mathbf{v}]_{1,1} \\ [\mathbf{v}]_{1,2} \\ [\mathbf{v}]_{2,1} \\ [\mathbf{v}]_{2,2} \end{bmatrix}.$$

Určete

a) matici přechodu $\mathsf{P}_{\mathcal{E} \leftarrow \mathcal{F}}$ od souřadnic v bázi $\mathcal{E}$ k souřadnicím v bázi $\mathcal{F}$,

b) matici $\mathsf{A}_{\mathcal{F} \leftarrow \mathcal{E}}$ lineárního zobrazení A vzhledem k bázím $\mathcal{E}$ a $\mathcal{F}$,

c) matici $\mathsf{B}_{\mathcal{E},\mathcal{F}}$ bilineární formy B vzhledem k bázím $\mathcal{E}$ a $\mathcal{F}$.

Řesení. □

a) Uvedeme si dva způsoby, jak $\mathsf{P}_{\mathcal{E} \leftarrow \mathcal{F}}$ spočíst.

i) přímo ze vzorce $\mathsf{P}_{\mathcal{F} \leftarrow \mathcal{E}} = \begin{bmatrix} [\mathbf{e}_1]_{\mathcal{F}} & \dots & [\mathbf{e}_4]_{\mathcal{F}} \end{bmatrix}$, kde pro zkrácení zápisu píšeme $[\mathbf{e}_i]_{\mathcal{F},j} = p_{j,i}$:

$$\mathbf{e}_i = p_{1,i}\mathbf{f}_1 + p_{2,i}\mathbf{f}_2 + p_{3,i}\mathbf{f}_3 + p_{4,i}\mathbf{f}_4 \quad i = 1, \dots, 4$$

$$\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} = p_{1,1} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + p_{2,1} \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix} + p_{3,1} \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix} + p_{4,1} \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$$

$$\vdots$$

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = p_{1,4} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + p_{2,4} \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix} + p_{3,4} \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix} + p_{4,4} \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}.$$

Tyto čtyři rovnosti mezi $\mathbb{R}^{2\times 2}$ maticemi jsou ekvivalentní s rovnostmi

$$\begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = p_{1,1} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} + p_{2,1} \begin{bmatrix} 4 \\ 1 \\ 2 \\ 3 \end{bmatrix} + p_{3,1} \begin{bmatrix} 3 \\ 4 \\ 1 \\ 2 \end{bmatrix} + p_{4,1} \begin{bmatrix} 2 \\ 3 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 3 & 2 \\ 2 & 1 & 4 & 3 \\ 3 & 2 & 1 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} p_{1,1} \\ p_{2,1} \\ p_{3,1} \\ p_{4,1} \end{bmatrix}$$

$$\vdots$$

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} = p_{1,4} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} + p_{2,4} \begin{bmatrix} 4 \\ 1 \\ 2 \\ 3 \end{bmatrix} + p_{3,4} \begin{bmatrix} 3 \\ 4 \\ 1 \\ 2 \end{bmatrix} + p_{4,4} \begin{bmatrix} 2 \\ 3 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 3 & 2 \\ 2 & 1 & 4 & 3 \\ 3 & 2 & 1 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} p_{1,4} \\ p_{2,4} \\ p_{3,4} \\ p_{4,4} \end{bmatrix},$$

což lze zapsat dohromady jako

$$\begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 3 & 2 \\ 2 & 1 & 4 & 3 \\ 3 & 2 & 1 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} & p_{1,4} \\ p_{2,1} & p_{2,2} & p_{2,3} & p_{2,4} \\ p_{3,1} & p_{3,2} & p_{3,3} & p_{3,4} \\ p_{4,1} & p_{4,2} & p_{4,3} & p_{4,4} \end{bmatrix}$$

$$\begin{aligned} \mathsf{P}_{\mathcal{F}\leftarrow\mathcal{E}} = \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} & p_{1,4} \\ p_{2,1} & p_{2,2} & p_{2,3} & p_{2,4} \\ p_{3,1} & p_{3,2} & p_{3,3} & p_{3,4} \\ p_{4,1} & p_{4,2} & p_{4,3} & p_{4,4} \end{bmatrix} &= \begin{bmatrix} 1 & 4 & 3 & 2 \\ 2 & 1 & 4 & 3 \\ 3 & 2 & 1 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix} \\ &= \frac{1}{40} \begin{bmatrix} -9 & 1 & 1 & 11 \\ 11 & -9 & 1 & 1 \\ 1 & 11 & -9 & 1 \\ 1 & 1 & 11 & -9 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix} \\ &= \frac{1}{20} \begin{bmatrix} 11 & 1 & 1 & -9 \\ -9 & 11 & 1 & 1 \\ 1 & -9 & 11 & 1 \\ 1 & 1 & -9 & 11 \end{bmatrix}. \end{aligned}$$

i) stejný výsledek dostaneme použitím standardní báze

$$\begin{aligned} \mathcal{S} &= \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right) \\ \mathsf{P}_{\mathcal{S}\leftarrow\mathcal{E}} &= \begin{bmatrix} [\mathbf{e}_1]_{\mathcal{S}} & \dots & [\mathbf{e}_4]_{\mathcal{S}} \end{bmatrix} = \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix} \\ \mathsf{P}_{\mathcal{S}\leftarrow\mathcal{F}} &= \begin{bmatrix} [\mathbf{f}_1]_{\mathcal{S}} & \dots & [\mathbf{f}_4]_{\mathcal{S}} \end{bmatrix} = \begin{bmatrix} 1 & 4 & 3 & 2 \\ 2 & 1 & 4 & 3 \\ 3 & 2 & 1 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix} \\ \mathsf{P}_{\mathcal{F}\leftarrow\mathcal{E}} &= \mathsf{P}_{\mathcal{F}\leftarrow\mathcal{S}} \mathsf{P}_{\mathcal{S}\leftarrow\mathcal{E}} = \mathsf{P}^{-1}_{\mathcal{S}\leftarrow\mathcal{F}} \mathsf{P}_{\mathcal{S}\leftarrow\mathcal{E}} \end{aligned}$$

protože $(\forall \mathbf{u} \in \mathbb{R}^{2\times 2})$:

$$[\mathbf{u}]_{\mathcal{F}} = \mathsf{P}_{\mathcal{F}\leftarrow\mathcal{S}}[\mathbf{u}]_{\mathcal{S}} = \mathsf{P}_{\mathcal{F}\leftarrow\mathcal{S}}\mathsf{P}_{\mathcal{S}\leftarrow\mathcal{E}}[\mathbf{u}]_{\mathcal{E}} = \mathsf{P}^{-1}_{\mathcal{S}\leftarrow\mathcal{F}}\mathsf{P}_{\mathcal{S}\leftarrow\mathcal{E}}[\mathbf{u}]_{\mathcal{E}} = \mathsf{P}_{\mathcal{F}\leftarrow\mathcal{E}}[\mathbf{u}]_{\mathcal{E}}\,.$$

b) Dle definice matice lineárního zobrazení je $(\forall \mathbf{u} \in \mathbb{R}^{2\times 2}) : [A(\mathbf{u})]_{\mathcal{S}} = \mathsf{A}_{\mathcal{S}\leftarrow\mathcal{S}}[\mathbf{u}]_{\mathcal{S}}$ a

$$A(\mathbf{s}_1) = \begin{bmatrix} 8 & 6 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 8 & 0 \\ 0 & 0 \end{bmatrix}, \ \ldots, \ A(\mathbf{s}_4) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}, \ \Rightarrow \ \mathsf{A}_{\mathcal{S}\leftarrow\mathcal{S}} = \begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

$$\begin{aligned}
\mathsf{P}_{\mathcal{S}\leftarrow\mathcal{F}}[A(\mathbf{u})]_{\mathcal{F}} &= [A(\mathbf{u})]_{\mathcal{S}} = \mathsf{A}_{\mathcal{S}\leftarrow\mathcal{S}}[\mathbf{u}]_{\mathcal{S}} = \mathsf{A}_{\mathcal{S}\leftarrow\mathcal{S}}\mathsf{P}_{\mathcal{S}\leftarrow\mathcal{E}}[\mathbf{u}]_{\mathcal{E}} \\
[A(\mathbf{u})]_{\mathcal{F}} &= \mathsf{A}_{\mathcal{E}\leftarrow\mathcal{F}}[\mathbf{u}]_{\mathcal{E}} = \mathsf{P}^{-1}_{\mathcal{S}\leftarrow\mathcal{F}}\mathsf{A}_{\mathcal{S}\leftarrow\mathcal{S}}\mathsf{P}_{\mathcal{S}\leftarrow\mathcal{E}}[\mathbf{u}]_{\mathcal{E}} \\
\mathsf{A}_{\mathcal{F}\leftarrow\mathcal{E}} &= \mathsf{P}^{-1}_{\mathcal{S}\leftarrow\mathcal{F}}\mathsf{A}_{\mathcal{S}\leftarrow\mathcal{S}}\mathsf{P}_{\mathcal{S}\leftarrow\mathcal{E}} = \begin{bmatrix} 1 & 4 & 3 & 2 \\ 2 & 1 & 4 & 3 \\ 3 & 2 & 1 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix} \\
&= \frac{1}{10}\begin{bmatrix} 26 & -13 & -12 & -21 \\ -34 & 37 & 8 & 9 \\ 6 & -23 & 28 & 9 \\ 6 & 7 & -12 & 19 \end{bmatrix}.
\end{aligned}$$

c) Připomeneme definici matice bilineární formy. Je-li $\tilde{B} : \mathcal{U} \times \mathcal{V} \to \mathbb{R}$ bilineární forma, $\mathcal{G}$ a $\mathcal{H}$ jsou báze prostorů $\mathcal{U}$ a $\mathcal{V}$, je matice $\tilde{\mathsf{B}}_{\mathcal{G},\mathcal{H}} \in \mathbb{R}^{\dim \mathcal{U} \times \dim \mathcal{V}}$ bilineární formy $\tilde{B}$ vzhledem k bázím $\mathcal{G}$ a $\mathcal{H}$ definována jako

$$\tilde{\mathsf{B}}_{\mathcal{G},\mathcal{H}} = \begin{bmatrix} \tilde{B}(\mathbf{g}_1, \mathbf{h}_1) & \ldots & \tilde{B}(\mathbf{g}_1, \mathbf{h}_{\dim \mathcal{V}}) \\ \vdots & \ddots & \vdots \\ \tilde{B}(\mathbf{g}_{\dim \mathcal{U}}, \mathbf{h}_1) & \ldots & \tilde{B}(\mathbf{g}_{\dim \mathcal{U}}, \mathbf{h}_{\dim \mathcal{V}}) \end{bmatrix}.$$

Je zřejmé, že matice $\mathsf{B}_{\mathcal{S},\mathcal{S}}$ bilineární formy B vzhledem ke standardní bázi $\mathcal{S}$ je

$$[\mathsf{B}_{\mathcal{S},\mathcal{S}}]_{i,j} = \begin{bmatrix} 11 & 12 & 13 & 14 \\ 21 & 22 & 23 & 24 \\ 31 & 32 & 33 & 34 \\ 41 & 42 & 43 & 44 \end{bmatrix}.$$

Využijme již napočítaných matic přechodu k získání matice $\mathsf{B}_{\mathcal{E},\mathcal{F}}$. Jelikož $(\forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^{2\times 2})$ je

$$\begin{aligned}
B(\mathbf{u}, \mathbf{v}) &= [\mathbf{u}]_{\mathcal{S}}^{\top}\mathsf{B}_{\mathcal{S},\mathcal{S}}[\mathbf{v}]_{\mathcal{S}} = (\mathsf{P}_{\mathcal{S}\leftarrow\mathcal{E}}[\mathbf{u}]_{\mathcal{E}})^{\top}\,\mathsf{B}_{\mathcal{S},\mathcal{S}}\mathsf{P}_{\mathcal{S}\leftarrow\mathcal{F}}[\mathbf{v}]_{\mathcal{F}} = [\mathbf{u}]_{\mathcal{E}}^{\top}\mathsf{P}_{\mathcal{S}\leftarrow\mathcal{E}}^{\top}\mathsf{B}_{\mathcal{S},\mathcal{S}}\mathsf{P}_{\mathcal{S}\leftarrow\mathcal{F}}[\mathbf{v}]_{\mathcal{F}} \\
\mathsf{B}_{\mathcal{E},\mathcal{F}} &= \mathsf{P}_{\mathcal{S}\leftarrow\mathcal{E}}^{\top}\mathsf{B}_{\mathcal{S},\mathcal{S}}\mathsf{P}_{\mathcal{S}\leftarrow\mathcal{F}} = \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}^{\top} \cdot \begin{bmatrix} 11 & 12 & 13 & 14 \\ 21 & 22 & 23 & 24 \\ 31 & 32 & 33 & 34 \\ 41 & 42 & 43 & 44 \end{bmatrix} \cdot \begin{bmatrix} 1 & 4 & 3 & 2 \\ 2 & 1 & 4 & 3 \\ 3 & 2 & 1 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 860 & 848 & 844 & 848 \\ 660 & 648 & 644 & 648 \\ 460 & 448 & 444 & 448 \\ 260 & 248 & 244 & 248 \end{bmatrix} = 8 \cdot \begin{bmatrix} 215 & 212 & 211 & 212 \\ 165 & 162 & 161 & 162 \\ 115 & 112 & 111 & 112 \\ 65 & 62 & 61 & 62 \end{bmatrix}.
\end{aligned}$$

Spočítejme matici $\mathsf{B}_{\mathcal{E},\mathcal{F}}$ pro kontrolu ještě přímo z definice

$$[\mathsf{B}_{\mathcal{E},\mathcal{F}}]_{1,1} = B(\mathbf{e}_1, \mathbf{f}_1) = \begin{bmatrix} -1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 11 & 12 & 13 & 14 \\ 21 & 22 & 23 & 24 \\ 31 & 32 & 33 & 34 \\ 41 & 42 & 43 & 44 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = 860$$

$$\vdots$$

$$[\mathsf{B}_{\mathcal{E},\mathcal{F}}]_{4,4} = B(\mathbf{e}_4, \mathbf{f}_4) = \begin{bmatrix} 1 & 1 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 11 & 12 & 13 & 14 \\ 21 & 22 & 23 & 24 \\ 31 & 32 & 33 & 34 \\ 41 & 42 & 43 & 44 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \\ 4 \\ 1 \end{bmatrix} = 248.$$

1.4 Skalární součin, norma, ortonormální báze, unitární matice

Bilineární formu $(\cdot,\cdot) : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ nazveme pojmem <u>skalární součin</u> právě tehdy, když je symetrická a pozitivně definitní, tedy $(\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V})(\forall \alpha \in \mathbb{R})$:

$$\begin{aligned} (\mathbf{u} \oplus \mathbf{v}, \mathbf{w}) &= (\mathbf{u}, \mathbf{w}) + (\mathbf{v}, \mathbf{w}) \\ (\alpha \odot \mathbf{u}, \mathbf{v}) &= \alpha(\mathbf{u}, \mathbf{v}) \\ (\mathbf{u}, \mathbf{v}) &= (\mathbf{v}, \mathbf{u}) \\ (\mathbf{u}, \mathbf{u}) &> 0 \text{ pro } \mathbf{u} \neq 0. \end{aligned}$$

Zobrazení $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ se nazývá <u>norma</u>, jestliže $(\forall \mathbf{u}, \mathbf{v} \in \mathcal{V})(\forall \alpha \in \mathbb{R})$:

$$\begin{aligned} \|\mathbf{u} \oplus \mathbf{v}\| &\leq \|\mathbf{u}\| + \|\mathbf{v}\| \\ \|\alpha \odot \mathbf{u}\| &= |\alpha| \, \|\mathbf{u}\| \\ \|\mathbf{u}\| = 0 \;&\Leftrightarrow\; \mathbf{u} = \mathbf{0}. \end{aligned}$$

Skalární součin $(\cdot,\cdot)_a$ definuje na $\mathcal{V}$ normu $\|\mathbf{u}\|_a = \sqrt{(\mathbf{u}, \mathbf{u})_a}$, ne každá norma je generována skalárním součinem.

Mějme vektorový prostor se skalárním součinem $(\mathcal{V}, F, \oplus, \odot, (\cdot,\cdot))$ a uspořádanou množinu vektorů $\mathcal{E} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$. Množinu $\mathcal{E}$ nazveme

$$\underline{\text{ortogonální}} \;\Leftrightarrow\; \begin{cases} (\mathbf{e}_i, \mathbf{e}_j) = 0 & i \neq j \\ (\mathbf{e}_i, \mathbf{e}_j) > 0 & i = j \end{cases}$$

$$\underline{\text{ortonormální}} \;\Leftrightarrow\; \begin{cases} (\mathbf{e}_i, \mathbf{e}_j) = 0 & i \neq j \\ (\mathbf{e}_i, \mathbf{e}_j) = 1 & i = j. \end{cases}$$

Příklad 6. Nechť $(\mathcal{V}, F, \oplus_{\mathcal{V}}, \odot_{\mathcal{V}}, (\cdot,\cdot))$ je vektorový prostor se skalárním součinem a vzhledem k němu ortonormální bází $\mathcal{E} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$. Dokažte, že souřadnice $[\mathbf{v}]_{\mathcal{E}}$ lze vypočítat pomocí vzorce $[\mathbf{v}]_{\mathcal{E},i} = (\mathbf{v}, \mathbf{e}_i)$. Výhodou tohoto postupu je možnost paralelizace.

Řesení. Souřadnice $[\mathbf{v}]_{\mathcal{E}}$ jsou dle definice koeficienty lineární kombinace

$$\mathbf{v} = [\mathbf{v}]_{\mathcal{E},1} \odot_{\mathcal{V}} \mathbf{e}_1 \oplus_{\mathcal{V}} \cdots \oplus_{\mathcal{V}} [\mathbf{v}]_{\mathcal{E},n} \odot_{\mathcal{V}} \mathbf{e}_n \,.$$

Proveďme skalární součin obou stran předchozí rovnosti s vektorem $\mathbf{e}_i$

$$\begin{aligned} (\mathbf{v}, \mathbf{e}_i) &= ([\mathbf{v}]_{\mathcal{E},1} \odot_{\mathcal{V}} \mathbf{e}_1 \oplus_{\mathcal{V}} \cdots \oplus_{\mathcal{V}} [\mathbf{v}]_{\mathcal{E},n} \odot_{\mathcal{V}} \mathbf{e}_n, \mathbf{e}_i) \\ &= [\mathbf{v}]_{\mathcal{E},1}(\mathbf{e}_1, \mathbf{e}_i) + \cdots + [\mathbf{v}]_{\mathcal{E},n}(\mathbf{e}_n, \mathbf{e}_i) = [\mathbf{v}]_{\mathcal{E},1}\delta_{1,i} + \cdots + [\mathbf{v}]_{\mathcal{E},n}\delta_{n,i} = [\mathbf{v}]_{\mathcal{E},i} \,. \end{aligned}$$

□

Příklad 7. Připomeňme si Grammův-Schmidtův ortonormalizační proces. Mějme ve vektorovém prostoru se skalárním součinem $(\mathbb{R}^4, \mathbb{R}, \oplus_{\mathbb{R}^4}, \odot_{\mathbb{R}^4}, (\cdot,\cdot)_b)$,

$$\left(\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}, \begin{bmatrix} e \\ f \\ g \\ h \end{bmatrix} \right)_b = \begin{bmatrix} a & b & c & d \end{bmatrix} \begin{bmatrix} 4 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} e \\ f \\ g \\ h \end{bmatrix}$$

podprostor

$$\langle \mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3 \rangle = \left\langle \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\rangle = \mathcal{V} \subsetneq \mathbb{R}^4.$$

Pomocí Grammova-Schmidtova ortonormalizačního procesu vytvořme ortonormální bázi $\mathcal{E}$ (vzhledem ke skalárnímu součinu $(\cdot,\cdot)_b$) prostoru $\mathcal{V}$.

Řesení. Grammovův-Schmidtův ortonormalizační proces je popsán v [1]. Postupně vytvoříme ortonormalizovanou bázi prostorů

$$\langle \mathbf{e}_1 \rangle = \langle \mathbf{f}_1 \rangle = \mathcal{V}_1, \quad \langle \mathbf{e}_1, \mathbf{e}_2 \rangle = \langle \mathbf{f}_1, \mathbf{f}_2 \rangle = \mathcal{V}_2, \quad \langle \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \rangle = \langle \mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3 \rangle = \mathcal{V}.$$

Pro hledanou ortonormální bázi $\mathcal{E} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ musí platit

$$\|\mathbf{e}_1\|_b = \sqrt{(\mathbf{e}_1, \mathbf{e}_2)_b} = 1, \quad \|\mathbf{e}_2\|_b = 1, \quad \|\mathbf{e}_3\|_b = 1, \quad (\mathbf{e}_1, \mathbf{e}_2)_b = 0, \quad (\mathbf{e}_1, \mathbf{e}_3)_b = 0, \quad (\mathbf{e}_2, \mathbf{e}_3)_b = 0.$$

1. Vektor $\mathbf{e}_1$ takový, že pro něj platí $\langle \mathbf{e}_1 \rangle = \langle \mathbf{f}_1 \rangle$ a současně $\|\mathbf{e}_1\|_b = 1$ je zřejmě

$$\mathbf{e}_1 = \frac{1}{\|\mathbf{f}_1\|_b} \odot_{\mathbb{R}^4} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{\begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 4 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}}} \odot_{\mathbb{R}^4} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{6}} \odot_{\mathbb{R}^4} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 6^{-\frac{1}{2}} \\ 6^{-\frac{1}{2}} \\ 0 \\ 0 \end{bmatrix}.$$

2. Zkonstruujme vektor $\tilde{\mathbf{e}}_2 \in \mathcal{V}_2 = \langle \mathbf{f}_1, \mathbf{f}_2 \rangle = \langle \mathbf{e}_1, \mathbf{f}_2 \rangle$, tedy $\tilde{\mathbf{e}}_2 = \mathbf{f}_2 \oplus_{\mathbb{R}^4} (-\alpha) \odot_{\mathbb{R}^4} \mathbf{e}_1$ tak aby

$$0 = (\tilde{\mathbf{e}}_2, \mathbf{e}_1)_b = (\mathbf{f}_2 \oplus_{\mathbb{R}^4} (-\alpha) \odot_{\mathbb{R}^4} \mathbf{e}_1, \mathbf{e}_1)_b = (\mathbf{f}_2, \mathbf{e}_1)_b - \alpha(\mathbf{e}_1, \mathbf{e}_1)_b = (\mathbf{f}_2, \mathbf{e}_1)_b - \alpha,$$

tedy

$$\alpha = (\mathbf{f}_2, \mathbf{e}_1)_b = \frac{1}{\sqrt{6}} \left(\begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 4 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right) = \frac{2}{\sqrt{6}}$$

$$\tilde{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \oplus_{\mathbb{R}^4} \left(-\frac{2}{\sqrt{6}} \right) \odot_{\mathbb{R}^4} \begin{bmatrix} 6^{-\frac{1}{2}} \\ 6^{-\frac{1}{2}} \\ 0 \\ 0 \end{bmatrix} = \frac{1}{3} \odot_{\mathbb{R}^4} \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix}$$

$$\mathbf{e}_2 = \frac{1}{\|\tilde{\mathbf{e}}_2\|_b} \odot_{\mathbb{R}^4} \tilde{\mathbf{e}}_2 = \frac{1}{\sqrt{\frac{1}{9} \left(\begin{bmatrix} -1 & 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 4 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix} \right)}} \frac{1}{3} \odot_{\mathbb{R}^4} \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix} = \frac{1}{4\sqrt{3}} \odot_{\mathbb{R}^4} \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix}.$$

3. Nyní vás stejně jako nás zřejmě otravuje neustálé používání zápisu $\oplus_{\mathbb{R}^4}$ a $\odot_{\mathbb{R}^4}$ při každém výskytu těchto operací. Proto je zde nebudeme explicitně zapisovat, jelikož jejich výskyt bude dostatečně zřejmý ze zápisu. Zkonstruujme vektor $\tilde{\mathbf{e}}_3 \in \mathcal{V} = \langle \mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3 \rangle = \langle \mathbf{e}_1, \mathbf{e}_2, \mathbf{f}_3 \rangle$, tedy $\tilde{\mathbf{e}}_3 = \mathbf{f}_3 - \alpha \mathbf{e}_1 - \beta \mathbf{e}_2$ tak aby

$$
\begin{aligned}
0 &= (\tilde{\mathbf{e}}_3, \mathbf{e}_1)_b = (\mathbf{f}_3 - \alpha \mathbf{e}_1 - \beta \mathbf{e}_2, \mathbf{e}_1)_b = (\mathbf{f}_3, \mathbf{e}_1)_b - \alpha(\mathbf{e}_1, \mathbf{e}_1)_b - \beta(\mathbf{e}_2, \mathbf{e}_1)_b = (\mathbf{f}_3, \mathbf{e}_1)_b - \alpha \\
0 &= (\tilde{\mathbf{e}}_3, \mathbf{e}_2)_b = (\mathbf{f}_3 - \alpha \mathbf{e}_1 - \beta \mathbf{e}_2, \mathbf{e}_2)_b = (\mathbf{f}_3, \mathbf{e}_2)_b - \alpha(\mathbf{e}_1, \mathbf{e}_2)_b - \beta(\mathbf{e}_2, \mathbf{e}_2)_b = (\mathbf{f}_3, \mathbf{e}_2)_b - \beta,
\end{aligned}
$$

tedy

$$
\begin{aligned}
\alpha = (\mathbf{f}_3, \mathbf{e}_1)_b &= \frac{1}{\sqrt{6}} \left(\begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right) = \frac{-1}{\sqrt{6}} \\
\beta = (\mathbf{f}_3, \mathbf{e}_2)_b &= \frac{1}{4\sqrt{3}} \left(\begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix} \right) = \frac{7}{4\sqrt{3}} \\
\tilde{\mathbf{e}}_2 &= \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \left(\frac{-1}{\sqrt{6}} \right) \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \left(\frac{7}{4\sqrt{3}} \right) \frac{1}{4\sqrt{3}} \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix} = \frac{1}{16} \begin{bmatrix} 5 \\ -2 \\ 9 \\ 16 \end{bmatrix} \\
\mathbf{e}_2 = \frac{1}{\|\tilde{\mathbf{e}}_3\|_b} \tilde{\mathbf{e}}_3 &= \frac{1}{\sqrt{\frac{1}{16^2} \left(\begin{bmatrix} 5 & -2 & 9 & 16 \end{bmatrix} \begin{bmatrix} 4 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ -2 \\ 9 \\ 16 \end{bmatrix} \right)}} \frac{1}{16} \begin{bmatrix} 5 \\ -2 \\ 9 \\ 16 \end{bmatrix} = \frac{1}{4\sqrt{77}} \begin{bmatrix} 5 \\ -2 \\ 9 \\ 16 \end{bmatrix}.
\end{aligned}
$$

Shrnuto

$$
\mathcal{E} = \left(\frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \frac{1}{4\sqrt{3}} \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix}, \frac{1}{4\sqrt{77}} \begin{bmatrix} 5 \\ -2 \\ 9 \\ 16 \end{bmatrix} \right).
$$

$\square$

2 Rychlá Fourierova transformace (FFT)

2.1 Komplexní čísla

K vektorovému prostoru $(\mathbb{R}^2, \mathbb{R}, \oplus_{\mathbb{R}^2}, \odot_{\mathbb{R}^2})$ přidáme operaci násobení vektorem

$$\begin{bmatrix} a \\ b \end{bmatrix} \odot_{\mathbb{R}^2,\mathbb{R}^2} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} ac - bd \\ ad + bc \end{bmatrix}$$

a označíme $(\mathbb{R}^2, \mathbb{R}, \oplus_{\mathbb{R}^2}, \odot_{\mathbb{R}^2}, \odot_{\mathbb{R}^2,\mathbb{R}^2})$ takto vzniklou pětici $\mathbb{C}$. Lze ukázat, že $(\mathbb{C}, \oplus_{\mathbb{C}}, \odot_{\mathbb{C}})$, kde

$$\mathbb{C} = \mathbb{R}^2, \quad \oplus_{\mathbb{C}} = \oplus_{\mathbb{R}^2}, \quad \odot_{\mathbb{C}} = \odot_{\mathbb{R}^2}$$

je těleso, které nazveme pojmem těleso komplexních čísel. Obvykle zapisujeme

$$a + bi = \begin{bmatrix} a \\ b \end{bmatrix}.$$

Více viz [2], kapitola 2.

Příklad 8. Ukažte, že inverzní komplexní číslo je nutno definovat jako

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \frac{a}{a^2+b^2} \\ \frac{-b}{a^2+b^2} \end{bmatrix}.$$

Řesení. Inverzní komplexních číslo z^{-1} k číslu $z = a+bi$ je třeba nadefinovat tak, aby platilo $z \odot_{\mathbb{C}} z^{-1} = w$, kde w je neutrální prvek vzhledem k násobení komplexních čísel, tj

$$z \odot_{\mathbb{C}} w = z \;\Rightarrow\; (a + bi) = (a + bi) \odot_{\mathbb{C}} (w_1 + w_2 i) = (aw_1 - bw_2 + i(aw_2 + bw_1))\,,$$

což platí pouze pro $w = (1 + 0i)$. Proto

$$z^{-1} = (c + di) : \; (1 + 0i) = w = z \odot_{\mathbb{C}} z^{-1} = (a + bi) \odot_{\mathbb{C}} (c + di) = (ac - bd + i(ad + bc))\,.$$

Rozepsáním rovnosti mezi komplexními čísly dostaneme

$$\begin{aligned} 1 &= ac - bd && (\mathrm{r}_1) \\ 0 &= bc + ad && (\mathrm{r}_2) \\ a\mathrm{r}_1 + b\mathrm{r}_2 : \; a &= (a^2 + b^2)c \\ b\mathrm{r}_1 - a\mathrm{r}_2 : \; b &= -(a^2 + b^2)d, \end{aligned}$$

což je ekvivalentní vzorci, který jsme chtěli dokázat. □

2.2 Pojmy z integrálních a diskrétních transformací

- Fourierova transformace

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) \mathrm{e}^{-2\pi i x \xi} \, \mathrm{d}x \qquad f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi) \mathrm{e}^{2\pi i \xi x} \, \mathrm{d}\xi$$

- Fourierovy řady

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) \, \mathrm{d}x \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) \, \mathrm{d}x \end{aligned}, \qquad f(x) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx)$$

- Diskrétní Fourierova transformace (DFT)

$$\hat{f}_n = \sum_{k=0}^{N-1} f_k \mathrm{e}^{-\frac{2\pi i n k}{N}}, \qquad f_n = \frac{1}{N}\sum_{k=0}^{N-1} f_k \mathrm{e}^{\frac{2\pi i n k}{N}}$$

2.3 (DFT) Diskrétní Fourierova transformace (v 1D)

Poznámka 9. V této kapitole budeme indexovat složky vektorů od nuly (jako v `C`, `C++`,...) a ne od jedničky (jako v `MatLabu`).

Mějme vektorový prostor se skalárním součinem $\left(\mathbb{C}^N, \mathbb{C}, \oplus, \odot, \langle\cdot,\cdot\rangle\right)$,

$$\langle \mathbf{u}, \mathbf{v}\rangle = \sum_{s=0}^{N-1} [\mathbf{u}]_s \overline{[\mathbf{v}]}_s\,,$$

a v něm bázi $\mathcal{E} = (\mathbf{e}_0, \dots, \mathbf{e}_{N-1})$,

$$\mathbf{e}_k = \frac{1}{\sqrt{N}} \begin{bmatrix} \mathrm{e}^{\frac{2\pi i k 0}{N}} & \dots & \mathrm{e}^{\frac{2\pi i k l}{N}} & \dots & \mathrm{e}^{\frac{2\pi i k (N-1)}{N}} \end{bmatrix}^\top, \quad k \in \{0, \dots, N-1\}.$$

Jelikož

$$\begin{aligned} \langle \mathbf{e}_k, \mathbf{e}_l\rangle &= \sum_{s=0}^{N-1} [\mathbf{e}_k]_s \overline{[\mathbf{e}_l]}_s = \sum_{s=0}^{N-1} \frac{1}{\sqrt{N}} \mathrm{e}^{\frac{2\pi i k s}{N}} \overline{\frac{1}{\sqrt{N}} \mathrm{e}^{\frac{2\pi i l s}{N}}} = \frac{1}{N} \sum_{s=0}^{N-1} \mathrm{e}^{\frac{2\pi i k s}{N}} \mathrm{e}^{\frac{-2\pi i l s}{N}} \\ &= \frac{1}{N} \sum_{s=0}^{N-1} \mathrm{e}^{\frac{2\pi i (k-l) s}{N}} = \frac{1}{N} \sum_{s=0}^{N-1} \left(\mathrm{e}^{\frac{2\pi i (k-l)}{N}}\right)^s = \begin{cases} \frac{1}{N} \cdot \frac{1-\mathrm{e}^{\frac{2\pi i (k-l) N}{N}}}{1-\mathrm{e}^{\frac{2\pi i (k-l)}{N}}} = 0 & k \neq l \\ \frac{N}{N} = 1 & k = l\,, \end{cases} \end{aligned}$$

je báze $\mathcal{E}$ ortonormální vzhledem ke skalárnímu součinu $\langle\cdot,\cdot\rangle$. Ve výpočtu jsme využili součtového vzorce pro geometrickou posloupnost

$$\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q}.$$

Tento jednoduchý postup (a zřejmě ne příliš efektivní) je naprogramován v `Matlab`ovském Kódu 5

Listing 5: dft_matrix.m

```
function DFT = dft_matrix(n)
ii   = 0:1:n-1;          % row vector for i
jj   = 0:1:n-1;          % row vector for j
w    = exp(-1j*2*pi/n);  % factor (w)
DFT  = w.^(ii'*jj);      % DFT matrix
```

Příklad 10. Zkonstruujte bázi $\mathcal{E}$ pro $\mathbb{C}^4$. Určete $\left[\begin{bmatrix}1 & 2 & 3 & 4\end{bmatrix}^\top\right]_{\mathcal{E}}$.

Řesení. Dle definice báze $\mathcal{E}$ je

$$\begin{aligned} \mathcal{E} = (\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) &= \left(\frac{1}{\sqrt{4}} \begin{bmatrix} \mathrm{e}^{\frac{2\pi i \cdot 0 \cdot 0}{4}} \\ \mathrm{e}^{\frac{2\pi i \cdot 0 \cdot 1}{4}} \\ \mathrm{e}^{\frac{2\pi i \cdot 0 \cdot 2}{4}} \\ \mathrm{e}^{\frac{2\pi i \cdot 0 \cdot 3}{4}} \end{bmatrix}, \frac{1}{\sqrt{4}} \begin{bmatrix} \mathrm{e}^{\frac{2\pi i \cdot 1 \cdot 0}{4}} \\ \mathrm{e}^{\frac{2\pi i \cdot 1 \cdot 1}{4}} \\ \mathrm{e}^{\frac{2\pi i \cdot 1 \cdot 2}{4}} \\ \mathrm{e}^{\frac{2\pi i \cdot 1 \cdot 3}{4}} \end{bmatrix}, \frac{1}{\sqrt{4}} \begin{bmatrix} \mathrm{e}^{\frac{2\pi i \cdot 2 \cdot 0}{4}} \\ \mathrm{e}^{\frac{2\pi i \cdot 2 \cdot 1}{4}} \\ \mathrm{e}^{\frac{2\pi i \cdot 2 \cdot 2}{4}} \\ \mathrm{e}^{\frac{2\pi i \cdot 2 \cdot 3}{4}} \end{bmatrix}, \frac{1}{\sqrt{4}} \begin{bmatrix} \mathrm{e}^{\frac{2\pi i \cdot 3 \cdot 0}{4}} \\ \mathrm{e}^{\frac{2\pi i \cdot 3 \cdot 1}{4}} \\ \mathrm{e}^{\frac{2\pi i \cdot 3 \cdot 2}{4}} \\ \mathrm{e}^{\frac{2\pi i \cdot 3 \cdot 3}{4}} \end{bmatrix} \right) \\ &= \left(\frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 1 \\ i \\ -1 \\ -i \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 1 \\ -i \\ -1 \\ i \end{bmatrix} \right). \end{aligned}$$

n	10	10^3	10^6	10^9	10^{12}
n^2	10^2	10^6	10^{12}	10^{18}	10^{24}
$n \log_2 n$	3.3e1	10^4	$2 \cdot 10^7$	$3 \cdot 10^{10}$	$4 \cdot 10^{13}$

Tabulka 2: Srovnání funkcí n^2 a $n \log_2 n$.

$\square$

Nyní bychom mohli využít ortonormální báze $\mathcal{E}$ pro zavedení diskrétní transformace přiřazující vektoru $\mathbf{u} = [\mathbf{u}]_{\mathcal{S}} \in \mathbb{C}^N$ jeho souřadnice $[\mathbf{u}]_{\mathcal{E}} \in \mathbb{C}^N$ v bázi $\mathcal{E}$. Toto lineární zobrazení by popisovaly matice transformace $\mathsf{P}_{\mathcal{E}\leftarrow\mathcal{S}}$ a zpětné transformace $\mathsf{P}_{\mathcal{S}\leftarrow\mathcal{E}} = \mathsf{P}_{\mathcal{E}\leftarrow\mathcal{S}}^{-1}$. Z historických důvodů je však diskrétní Fourierova transformace zavedena jako lineární zobrazení přiřazující vektoru $\mathbf{u} = [\mathbf{u}]_{\mathcal{S}} \in \mathbb{C}^N$ vektor $\hat{\mathbf{u}} \overset{\text{ozn.}}{\equiv} \sqrt{N}[\mathbf{u}]_{\mathcal{E}} \in \mathbb{C}^N$. Jelikož

$$
\begin{aligned}
\mathbf{u} &= \sum_{l=0}^{N-1} [\mathbf{u}]_{\mathcal{E},l} \mathbf{e}_l \\
\langle \mathbf{u}, \mathbf{e}_k \rangle &= \sum_{l=0}^{N-1} [\mathbf{u}]_{\mathcal{E},l} \langle \mathbf{e}_l, \mathbf{e}_k \rangle = \sum_{l=0}^{N-1} [\mathbf{u}]_{\mathcal{E},l} \delta_{k,l} = [\mathbf{u}]_{\mathcal{E},k} = \frac{1}{\sqrt{N}} \hat{\mathbf{u}}_k \,,
\end{aligned}
$$

platí pro diskrétní Fourierův obraz

$$
\hat{\mathbf{u}}_k = \sqrt{N}[\mathbf{u}]_{\mathcal{E},k} = \sqrt{N}\langle \mathbf{u}, \mathbf{e}_k \rangle = \sqrt{N} \sum_{n=0}^{N-1} [\mathbf{u}]_n \overline{[\mathbf{e}_k]}_n = \sum_{n=0}^{N-1} \mathbf{u}_n \mathrm{e}^{\frac{-2\pi i k n}{N}} . \tag{11}
$$

Dle tohoto vzorce je pro výpočet $\hat{\mathbf{u}}$ nutné určit všechny (je jich N) složky a výpočet každé z nich zahrnuje provedení N komplexních násobení a $N-1$ komplexních sčítání. Potřebujeme tedy cca $O(N^2)$ operací.

Z obrazu $\hat{\mathbf{u}}$ lze zpětně rekonstruovat původní $\mathbf{u}$ pomocí zpětné Fourierovy transformace

$$
\mathbf{u}_k = \frac{1}{N} \sum_{n=0}^{N-1} \hat{\mathbf{u}}_n \mathrm{e}^{\frac{2\pi i k n}{N}} . \tag{12}
$$

2.4 (FFT) Rychlá Fourierova transformace (v 1D)

Uvedeme si postup, který potřebuje pouze $O(N \log_2 N)$ operací, což pro velká N umožňuje několikařádově rychlejší transformaci (a také zpětnou transformaci, jelikož ta je výpočetně podobná), viz tabulka 2. Pro jednoduchost předpokládejme, že $N = 2^b$. Označme

$$
w_N = \mathrm{e}^{\frac{-2\pi i}{N}} \tag{13}
$$

a uvědomme si, že pro w_N platí

$$
\begin{aligned}
\text{(symetrie):} \quad & w_N^{k(N-n)} = w_N^{-kn} = \overline{w_N^{kn}} \\
\text{(periodicita):} \quad & w_N^{kn} = w_N^{k(N+n)} = w_N^{(k+N)n} .
\end{aligned}
$$

Pak $(\forall k \in 0, \dots, N-1)$

$$
\begin{aligned}
\hat{\mathbf{u}}_k &= \sum_{n=0}^{N-1} \mathbf{u}_n \mathrm{e}^{\frac{-2\pi i k n}{N}} = \sum_{\substack{n \in 0,\dots,N-1 \\ n \text{ je sudé}}} \mathbf{u}_n w_N^{kn} + \sum_{\substack{n \in 0,\dots,N-1 \\ n \text{ je liché}}} \mathbf{u}_n w_N^{kn} \\
&= \sum_{r=0}^{\frac{N}{2}-1} \mathbf{u}_{2r} w_N^{k2r} + \sum_{r=0}^{\frac{N}{2}-1} \mathbf{u}_{2r+1} w_N^{k(2r+1)} = \sum_{r=0}^{\frac{N}{2}-1} \mathbf{u}_{2r} \left(w_N^2\right)^{kr} + w_N^k \sum_{r=0}^{\frac{N}{2}-1} \mathbf{u}_{2r+1} \left(w_N^2\right)^{kr} \\
&= \sum_{r=0}^{\frac{N}{2}-1} \mathbf{u}_{2r} w_{\frac{N}{2}}^{kr} + w_N^k \sum_{r=0}^{\frac{N}{2}-1} \mathbf{u}_{2r+1} w_{\frac{N}{2}}^{kr} \\
&= \hat{\mathbf{u}}_{\text{sudé},k} + w_N^k \hat{\mathbf{u}}_{\text{liché},k} \,,
\end{aligned}
$$

kde jsme v poslední rovnosti využili periodicity $w_{\frac{N}{2}}^{rk} = w_{\frac{N}{2}}^{r(k+\frac{N}{2})}$. Rozepsáním poslední rovnosti dostaneme $(\forall k \in 0, \dots, \frac{N}{2}-1)$

$$
\begin{aligned}
\hat{\mathbf{u}}_k &= \hat{\mathbf{u}}_{\text{sudé},k} + w_N^k \hat{\mathbf{u}}_{\text{liché},k} \\
\hat{\mathbf{u}}_{k+\frac{N}{2}} &= \hat{\mathbf{u}}_{\text{sudé},k} + w_N^{k+\frac{N}{2}} \hat{\mathbf{u}}_{\text{liché},k} = \hat{\mathbf{u}}_{\text{sudé},k} - w_N^k \hat{\mathbf{u}}_{\text{liché},k}
\end{aligned}
$$

jelikož

$$
w_N^{k+\frac{N}{2}} = w_N^k w_N^{\frac{N}{2}} = w_N^k \mathrm{e}^{\frac{-2\pi i \frac{N}{2}}{N}} = w_N^k \mathrm{e}^{-\pi i} = -w_N^k \,.
$$

Tento jednoduchý postup je naprogramován v `Matlab`ovském Kódu 6 a testovací skript je uveden v Kódu 7.

Listing 6: fft_rec.m

```
function y = fft_rec(x)
if size(x,1)<size(x,1)
  x=x';
end
N = size(x,1);
if N == 1
  y = x;
else
  omega = exp(-2*pi*1i/N);
  k = (0:N/2-1)';
  w = omega.^k;
  y_e =    fft_rec( x(1:2:end-1) );  % FFT of even x
  y_o = w.*fft_rec( x(2:2:end   ) );  % FFT of odd  x
  y = [y_e+y_o; y_e-y_o];
end
```

Listing 7: fft_test.m

```
n = 11; eeps = 1e-3;
x = (0:2^n-1)'/(2^n-1);
y = 40*sin(1*2*pi*x)+30*cos(2*2*pi*x)+20*sin(3*2*pi*x)+10*cos(4*2*pi*x);
fprintf(1,'  Y:                              nnz: %03d\n',nnz(y));
% DFT
DFT = dft_matrix(2^n);
t_start = tic;     y_dft = DFT*y;          t_end = toc(t_start);
y_dft(abs(y_dft)<eeps*max(abs(y_dft))) = 0;          % tresholding
fprintf(1,'DFT : ... [%5.4f s], tresholded nnz: %03d\n',t_end,nnz(y_dft));
% FFT_REC
```

```
t_start = tic;     y_fftr = fft_rec(y);    t_end = toc(t_start);
y_fftr(abs(y_fftr)<eeps*max(abs(y_fftr))) = 0;      % tresholding
fprintf(1,'FFTr: ... [%5.4f s], tresholded nnz: %03d\n',t_end,nnz(y_fftr));
% FFT_ITER
t_start = tic;     y_ffti = fft_iter(y);   t_end = toc(t_start);
y_ffti(abs(y_ffti)<eeps*max(abs(y_ffti))) = 0;      % tresholding
fprintf(1,'FFTi: ... [%5.4f s], tresholded nnz: %03d\n',t_end,nnz(y_ffti));
% plotting
figure(1); hold on;
plot(x,y ,'-r');
plot(x,real(DFT'*y_dft )/2^n,'-b');
plot(x,real(DFT'*y_fftr)/2^n,'-k');
figure(2); hold on;
plot(x,real(y_dft),'-r');
```

Jelikož zpětná diskrétní Fourierova transformace (12) se od Fourierovy transformace (11) liší pouze znaménkem ve w_N (srovnej (13), (12) a (11)), lze algoritmem pro FFT počítat také zpětnou diskrétní Fourierovu transformaci (se dvěmi drobnými změnami – změnou znaménka v exponentu w_N a násobení N^{-1}). Uvedeme si ještě Kód 8, který je iteračním ekvivalentem rekurzivního kódu FFT.

Listing 8: fft_iter.m

```
function data_ = fft_iter(data_)

N = size(data_,1);
b = log2(N);
% preprocessing
for iii = 1:b-1
  cell_size  = 2^(b-iii+1);
  cell_count = 2^(iii-1);
  for jjj = 1:cell_count
    cell_ind = (jjj-1)*cell_size+1:jjj*cell_size;
    data_halfind_odd    = data_(cell_ind(1:2    :end-1));
    data_halfind_even   = data_(cell_ind(2:2    :end  ));
    cell_halfind_first  = cell_ind(1      :end/2);
    cell_halfind_second = cell_ind(end/2+1:end  );
    data_(cell_halfind_first)  = data_halfind_odd;
    data_(cell_halfind_second) = data_halfind_even;
  end
end
% iterative fft
for iii = 1:b
  cell_size = 2^iii;
  omega = exp(2*pi*1i/cell_size);
  w = (omega.^(0:cell_size/2-1))';
  for jjj = 1:2^(b-iii)
    cell_ind  = (jjj-1)*cell_size+1:jjj*cell_size;
    data_cell = data_(cell_ind);
    data_halfcell_first  =    data_cell(      1:end/2);
    data_halfcell_second = w.*data_cell(end/2+1:end);
    data_(cell_ind(1      :end/2)) = data_halfcell_first + data_halfcell_second;
    data_(cell_ind(end/2+1:end  )) = data_halfcell_first - data_halfcell_second;
  end
end
```

2.5 DFT ve 2D (ve více dimenzích)

Mějme vektorový prostor $(\mathbb{C}^{m\times n}, \mathbb{C},)$

$$\hat{\mathbf{u}}_{k,l} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \mathbf{u}_{m,n} \mathrm{e}^{-2\pi i\left(\frac{km}{M}+\frac{ln}{N}\right)}$$

$$\mathbf{u}_{m,n} = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \hat{\mathbf{u}}_{m,n} \mathrm{e}^{2\pi i\left(\frac{km}{M}+\frac{ln}{N}\right)}$$

3 Vybrané rozklady matic

3.1 Spektrální rozklad

Nechť $\mathsf{A} \in \mathbb{R}^{n\times n}$ je čtvercová matice. Vektor $\mathbf{v}$ nazveme vlastní vektor matice A právě tehdy když pro něj existuje skalár λ takový, že

$$\mathsf{A}\mathbf{v} = \lambda\mathbf{v}, \quad \text{což je totéž jako } (\mathsf{A} - \lambda\mathsf{I})\,\mathbf{v} = \mathbf{0}.$$

Geometricky to znamená, že vektory $\mathbf{v}$ a $\mathsf{A}\mathbf{v}$ jsou rovnoběžné. Dvojici $(\lambda, \mathbf{v})$ říkáme vlastní číslo a vlastní vektor matce A.

Matici $\mathsf{A} \in \mathbb{R}^{n\times n}$ s n lineárně nezávislými vlastními vektory lze zapsat ve tvaru

$$\mathsf{A} = \mathsf{Q}\Lambda\mathsf{Q}^{-1},$$

kde $\mathsf{Q} \in \mathbb{R}^{n\times n}$ je matice, která vznikne zapsáním n normalizovaných vlastních vektorů postupně do sloupců a matice $\Lambda \in \mathbb{R}^{n\times n}$ je diagonální a na její diagonále jsou postupně vlastní čísla matice A. Tomuto zápisu se říká vlastní rozklad matice A. Matice, které lze takto rozložit nazýváme diagonizovatelné, některé diagonalizovat nejdou.

Příklad 11. Spočítejte vlastní rozklad matic

$$\mathsf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix} \quad \text{a } \mathsf{B} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Řesení. Z definice spočteme vlastní čísla a vektory

- A:

$$\det(\mathsf{A} - \lambda\mathsf{I}) = \begin{vmatrix} 1-\lambda & 0 \\ 1 & 3-\lambda \end{vmatrix} = (1-\lambda)(3-\lambda) \;\Rightarrow\; \lambda \in \{1,3\}$$

$$\lambda_1 = 1: \quad \mathbf{0} = (\mathsf{A} - \lambda_1\mathsf{I})\,\tilde{\mathbf{v}}_1 = \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix}\tilde{\mathbf{v}}_1 \;\Rightarrow\; \tilde{\mathbf{v}}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix} p \;\Rightarrow\; \mathbf{v}_1 = \frac{\tilde{\mathbf{v}}_1}{\|\tilde{\mathbf{v}}_1\|} = \frac{1}{\sqrt{5}}\begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$\lambda_2 = 3: \quad \mathbf{0} = (\mathsf{A} - \lambda_2\mathsf{I})\,\tilde{\mathbf{v}}_2 = \begin{bmatrix} -2 & 0 \\ 1 & 0 \end{bmatrix}\tilde{\mathbf{v}}_2 \;\Rightarrow\; \tilde{\mathbf{v}}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} p \;\Rightarrow\; \mathbf{v}_2 = \frac{\tilde{\mathbf{v}}_2}{\|\tilde{\mathbf{v}}_2\|} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

  tedy

$$\mathsf{A} = \mathsf{QDQ}^{-1} = \begin{bmatrix} \frac{-2}{\sqrt{5}} & 0 \\ \frac{1}{\sqrt{5}} & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} \frac{-\sqrt{5}}{2} & 0 \\ \frac{1}{2} & 1 \end{bmatrix}.$$

□

- B:

$$\det(\mathsf{B} - \lambda\mathsf{I}) = \begin{vmatrix} 1-\lambda & 0 \\ 1 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 \;\Rightarrow\; \lambda_1 = \lambda_2 = 1$$

$$\mathbf{0} = (\mathsf{A} - \lambda\mathsf{I})\,\tilde{\mathbf{v}} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\tilde{\mathbf{v}} \;\Rightarrow\; \tilde{\mathbf{v}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} p$$

  vidíme, že k dvojnásobnému vlastnímu číslu $\lambda = 1$ lze v našem případě sestavit pouze jeden vlastní vektor $\mathbf{v} = [0\,1]^\top$. Chybí nám druhý vlastní vektor k sestavení matice Q. Matice A není diagonalizovatelná.

Co jsou vlastní čísla a vektory matice A^n?

Řesení. Z defininice

$$\mathsf{A}\mathbf{v} = \lambda\mathbf{v} \;\Rightarrow\; \mathsf{A}^n\mathbf{v} = \mathsf{A}^{n-1}\mathsf{A}\mathbf{v} = \mathsf{A}^{n-1}\lambda\mathbf{v} = \lambda\mathsf{A}^{n-1}\mathbf{v} = \lambda\mathsf{A}^{n-2}\mathsf{A}\mathbf{v} = \lambda^2\mathsf{A}^{n-2}\mathbf{v} = \lambda^n\mathbf{v},$$

což lze shrnout do závěru, že s mocninou A se vlastní vektory nemění, vlastní čísla se mění s mocninou. □

Dále se zabývejme některými vlastnostmi vlastního rozkladu

- Charakteristický polynom

  Pro matici $\mathsf{A} \in \mathbb{R}^{n\times n}$ je determinant $|\mathsf{A} - \lambda\mathsf{I}|$ je polynomem n–tého řádu proměnné λ. Tento polynom nazýváme charakteristický polynom. Dle základní věty algebry jej lze rozvést ve tvaru součinu kořenových činitelů

$$|\mathsf{A} - \lambda\mathsf{I}| = (\lambda_1 - \lambda)\cdots(\lambda_n - \lambda) = (-\lambda)^n + (\sum_{i=1}^{n}\lambda_i)(-\lambda)^{n-1} + \cdots + \prod_{i=1}^{n}\lambda_i. \tag{14}$$

- Invariantnost vzhledem k podobnosti matic

  Pokud je $(\lambda, \mathbf{v})$ vlastní dvojice matice A, tedy $\mathsf{A}\mathbf{v} = \lambda\mathbf{v}$. Pro jakoukoli matici B podobnou matici A, tedy $\mathsf{B} = \mathsf{TAT}^{-1}$, pak platí

$$\mathsf{BT}\mathbf{v} = \mathsf{TAT}^{-1}\mathsf{T}\mathbf{v} = \mathsf{TA}\mathbf{v} = \lambda\mathsf{T}\mathbf{v}.$$

  Platí tedy, že podobné matice mají stejná vlastní čísla. Navíc mají podobné matice také stejný charakteristický polynom, jelikož

$$\left|\mathsf{TAT}^{-1} - \lambda\mathsf{I}\right| = \left|\mathsf{TAT}^{-1} - \lambda\mathsf{TT}^{-1}\right| = \left|\mathsf{T}\left(\mathsf{A} - \lambda\mathsf{I}\right)\mathsf{T}^{-1}\right| = \left|\mathsf{T}\right|\left|\mathsf{A} - \lambda\mathsf{I}\right|\left|\mathsf{T}^{-1}\right| = \left|\mathsf{A} - \lambda\mathsf{I}\right|.$$

- Součin vlastních čísel matice

  Dosazením $\lambda = 0$ do charakteristického polynomu (14) dostaneme, že determinant A je roven součinu vlastních čísel.

- Součet vlastních čísel matice

  Charakteristický polynom lze také s využitím rozvoje determinantu podle řádku zapsat ve tvaru

$$|\mathsf{A} - \lambda\mathsf{I}| = \prod_{i=1}^{n}(\mathsf{A}_{ii} - \lambda) + p_{n-2}(\lambda) = p^{\star}_{n-2}(\lambda) + (\sum_{i=1}^{n}\mathsf{A}_{ii})(-\lambda)^{n-1} + (-\lambda)^n.$$

  Porovnáním koeficientů u $(-\lambda)^{n-1}$ s rozvojem (14) plyne, že $\sum_{i=1}^{n}\lambda_i = \sum_{i=1}^{n}\mathsf{A}_{ii}$.

Dále se budeme zabývat případem, kdy je matice A reálná a symetrická. V tomto případě sou vlastní čísla reálná, vlastní vektory příslušné vlastním číslům jsou ortogonální a vlastní vektory příslušné vícenásobnému vlastnímu číslu leží tvoří podprostor dimenze rovné této násobnosti. Vlastní vektory lze tedy zvolit ortonormální a z vlastního rozkladu matice se stane spektrální rozklad matice

$$\mathsf{A} = \mathsf{QDQ}^\top.$$

3.2 QR rozklad, transformační matice Householderových zrcadlení

Nechť $\mathsf{A} \in \mathbb{R}^{m \times n}$ je libovolná matice. Pak existují ortonormální matice $\mathsf{Q} \in \mathbb{R}^{m \times m}$ a horní trojúhelníková matice $\mathsf{R} \in \mathbb{R}^{m \times n}$ takové, že

$$\mathsf{A} = \mathsf{QR}.$$

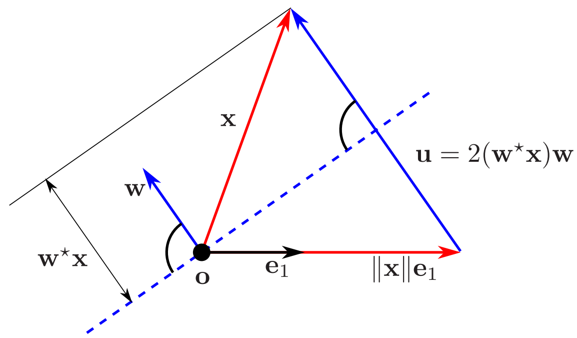
Probereme si jeden ze způsobů výpočtu QR rozkladu a to pomocí Householderových transformačních matic

$$\mathsf{P} = \mathsf{I} - 2\mathbf{w}\mathbf{w}^\star, \quad \|\mathbf{w}\| = 1,$$

které popisují lineární zobrazení vektoru $\mathbf{y}$ na vektor

$$\mathsf{P}\mathbf{y} = (\mathsf{I} - 2\mathbf{w}\mathbf{w}^\star)\,\mathbf{y} = \mathbf{y} - 2\mathbf{w}(\mathbf{w}^\star\mathbf{y}),$$

jež je jeho zrcadlovým obrazem, viz Obrázek 7. Toto zrcadlení je dáno vektorem $\mathbf{w}$, který je kolmý k nadrovině, podle které se zrcadlení provádí. Nejprve si dokážeme několik vlastností matice P.



Obrázek 7: Householderovo zrcadlení

- P je symetrická, jelikož $\mathsf{P}^\star = (\mathsf{I} - 2\mathbf{w}\mathbf{w}^\star)^\star = \mathsf{I}^\star - 2\,(\mathbf{w}\mathbf{w}^\star)^\star = \mathsf{I} - 2\,(\mathbf{w}^\star)^\star\,\mathbf{w}^\star = \mathsf{I} - 2\mathbf{w}\mathbf{w}^\star = \mathsf{P}$.
- P je ortogonální ($\mathsf{P}^{-1} = \mathsf{P}$), jelikož $\mathsf{PP} = (\mathsf{I} - 2\mathbf{w}\mathbf{w}^\star)\,(\mathsf{I} - 2\mathbf{w}\mathbf{w}^\star) = \mathsf{I} - 4\mathbf{w}\mathbf{w}^\star + 4\mathbf{w}\,\overbrace{\mathbf{w}^\star\mathbf{w}}^{=1}\,\mathbf{w}^\star = \mathsf{I}$.

Položme si otázku, jak získat rozklad matice $\mathsf{A} = \mathsf{QR}$. Výše jsme si ukázali jeden ze způsobů jak utvořit ortonormální matici P. Součin ortogonálních matic $\mathsf{P} = \mathsf{P}_n \cdots \mathsf{P}_2\mathsf{P}_1$ je ortonormální matice. Co kdybychom volili matice P_i tak, aby postupně převedli matici A na horní trojúhelníkovou matici R? Ideální by bylo, aby měla matice $\mathsf{P}_1\mathsf{A}$ v prvním sloupci nuly pod hlavní diagonálou, matice $\mathsf{P}_2\mathsf{P}_1\mathsf{A}$ by měla nuly pod diagonálou v prvním a druhém sloupci, a tak dále. Zbývá se přesvědčit, zdali lze pro matici A nalézt takový vektor $\mathbf{w}$, aby odpovídal zrcadlení, které převede první sloupec $\mathbf{x} = s_1(\mathsf{A})$ matice A na vektor $\|\mathbf{x}\|\mathbf{e}_1$ rovnoběžný s $\mathbf{e}_1$. Tento geometrický úkol jsme si ale již nastínili na Obrázku 7, kde

$$\mathsf{P}\mathbf{x} = \|\mathbf{x}\|\mathbf{e}_1 = \mathbf{x} - \mathbf{u}, \quad \mathbf{u} = \mathbf{x} - \mathsf{P}\mathbf{x}, \quad \mathbf{w} = \frac{\mathbf{u}}{\|\mathbf{u}\|}.$$

Takto lze provést QR rozklad matice A, existují však také efektivnější metody. Tělo QR faktorizace pomocí Householderových transformací je uvedeno v Kódu 10 a testováno souborem 9.

3.3 Singulární rozklad (SVD), transformační matice Givensových rotací

Nechť $\mathsf{A} \in \mathbb{R}^{m\times n}$ je libovolná matice. Pak existují ortonormální matice $\mathsf{U} \in \mathbb{R}^{m\times m}$ a $\mathsf{V} \in \mathbb{R}^{n\times n}$ a diagonální matice $\mathsf{S} \in \mathbb{R}^{m\times n}$, na jejíž diagonále jsou vlastní čísla matice

$$\sqrt{\mathsf{A}^\star\mathsf{A}} \text{ pro } m \geq n$$
$$\sqrt{\mathsf{A}\mathsf{A}^\star} \text{ pro } m < n$$

takové, že

$$\mathsf{A} = \mathsf{U}\mathsf{S}\mathsf{V}^\top.$$

Tomuto rozkladu říkáme <u>singulární rozklad</u>. Čísla na diagonále matice S se nazývají <u>singulární čísla</u> matice A.

- standadní verze (full SVD)

  $$\underset{m\times n}{\mathsf{A}} = \underset{m\times m}{\mathsf{U}}\underset{m\times n}{\mathsf{S}}\underset{n\times n}{\mathsf{V}^\top}$$

  v MatLabu: `[U,S,V]=svd(A)`

- redukovaná verze (reduced SVD)

  $$\underset{m\times n}{\mathsf{A}} = \begin{cases} \underset{m\times n}{\mathsf{U}}\underset{n\times n}{\mathsf{S}}\underset{n\times n}{\mathsf{V}^\star} & m > n \\ \underset{m\times m}{\mathsf{U}}\underset{m\times m}{\mathsf{S}}\underset{m\times n}{\mathsf{V}^\star} & m < n \end{cases}$$

  v MatLabu: `[U,S,V]=svd(A,'econ')`

- pro řídké matice v Matlabu (aproximace hodnosti q): `[U,S,V]=svds(A,q)`

- Aproximace matice maticí nižší hodnosti (low rank approximation). Nechť $h(\mathsf{A}) = r$ pak

  $$\mathsf{A} = \sum_{i=1}^{r} \mathsf{A}_i = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^\star$$

  Zkrácením této sumy z r na q členů dostaneme matici

  $$\hat{\mathsf{A}}_q = \sum_{i=1}^{q} \sigma_i \mathbf{u}_i \mathbf{v}_i^\star$$

  pro kterou platí

  $$\left\|\mathsf{A} - \hat{\mathsf{A}}_q\right\|_2 = \sigma_{q+1} \quad \text{kde } \|\mathsf{M}\|_2 \leq \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2} = \|\mathsf{M}\|_F$$

Příklad 12. Spočítejte singulární rozklad matice

$$\mathsf{A} = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} \quad \text{a } \mathsf{B} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Řesení. Z definice spočteme vlastní čísla a vektory

- A: Chceme najít vektory $\mathbf{v}_1, \mathbf{v}_1 \in \mathbb{R}^2$ a $\mathbf{u}_1, \mathbf{u}_1 \in \mathbb{R}^2$ a kladná čísla σ_1 a σ_2 taková, že $\mathsf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i$, tedy
$$\mathsf{AV} = \mathsf{US} \quad \Leftrightarrow \quad \mathsf{A} = \mathsf{USV}^\star.$$
Místo hledání trojice U, S, V budeme hledat pouze dvojici S, V:
$$\mathsf{A}^\top\mathsf{A} = \mathsf{VS}^\top\mathsf{U}^\top\mathsf{USV}^\top = \mathsf{VS}^2\mathsf{V}^\top = \mathsf{V}\mathrm{diag}(\sigma_1^2, \sigma_2^2)\mathsf{V}^\top$$
Tedy
$$\mathsf{A}^\top\mathsf{A} = \begin{bmatrix} 4 & -3 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 7 \\ 7 & 25 \end{bmatrix}$$
$$0 = \begin{vmatrix} 25-\lambda & 7 \\ 7 & 25-\lambda \end{vmatrix} = (25-\lambda)^2 - 49 = \lambda^2 - 50\lambda - 576 = (\lambda - 32)(\lambda - 18)$$
$$\begin{bmatrix} 25-32 & 7 \\ 7 & 25-32 \end{bmatrix} = \begin{bmatrix} -7 & 7 \\ 7 & -7 \end{bmatrix} \Rightarrow \mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
$$\begin{bmatrix} 25-18 & 7 \\ 7 & 25-18 \end{bmatrix} = \begin{bmatrix} 7 & 7 \\ 7 & 7 \end{bmatrix} \Rightarrow \mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$
$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{32} = 4\sqrt{2}$$
$$\sigma_2 = \sqrt{\lambda_2} = \sqrt{18} = 3\sqrt{2}$$
$$\mathsf{AV} = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \mathsf{U}\begin{bmatrix} 4\sqrt{2} & 0 \\ 0 & 2\sqrt{2} \end{bmatrix} = \mathsf{US}$$
$$\mathsf{U} = \mathsf{AVS}^{-1} = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 4^{-1} & 0 \\ 0 & 2^{-1} \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 0 & 8 \\ -3 & 0 \end{bmatrix}$$
$$\mathsf{A} = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 0 & 8 \\ -3 & 0 \end{bmatrix} \begin{bmatrix} 4\sqrt{2} & 0 \\ 0 & 2\sqrt{2} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \mathsf{USV}^\star$$
Jiný způsob: Vypočítáme spektrální rozklad symetrické matice
$$\begin{bmatrix} 0 & \mathsf{M} \\ \mathsf{M}^\star & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 4 & 4 \\ 0 & 0 & -3 & 3 \\ 4 & -3 & 0 & 0 \\ 4 & 3 & 0 & 0 \end{bmatrix}$$
$$0 = \begin{vmatrix} 0-\lambda & 0 & 4 & 4 \\ 0 & 0-\lambda & -3 & 3 \\ 4 & -3 & 0-\lambda & 0 \\ 4 & 3 & 0 & 0-\lambda \end{vmatrix} = \lambda^4 - 50\lambda^2 + 576 = (\lambda^2 - 32)(\lambda^2 - 18)$$
zkráceně
$$\lambda_1 = 4\sqrt{2}, \quad \mathbf{v}_1 = \frac{1}{2}\begin{bmatrix} \sqrt{2} \\ 0 \\ 1 \\ 1 \end{bmatrix}, \qquad \lambda_2 = 3\sqrt{2}, \quad \mathbf{v}_2 = \frac{1}{2}\begin{bmatrix} 0 \\ \sqrt{2} \\ -1 \\ 1 \end{bmatrix}$$
$$\lambda_3 = -4\sqrt{2}, \quad \mathbf{v}_3 = \frac{1}{2}\begin{bmatrix} -\sqrt{2} \\ 0 \\ 1 \\ 1 \end{bmatrix}, \qquad \lambda_4 = -3\sqrt{2}, \quad \mathbf{v}_4 = \frac{1}{2}\begin{bmatrix} 0 \\ -\sqrt{2} \\ -1 \\ 1 \end{bmatrix}$$

$$
\begin{bmatrix} 0 & \mathsf{M} \\ \mathsf{M}^\star & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 4 & 4 \\ 0 & 0 & -3 & 3 \\ 4 & -3 & 0 & 0 \\ 4 & 3 & 0 & 0 \end{bmatrix} =
$$

$$
= \frac{1}{2}\begin{bmatrix} \sqrt{2} & 0 & -\sqrt{2} & 0 \\ 0 & \sqrt{2} & 0 & -\sqrt{2} \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4\sqrt{2} & 0 & 0 & 0 \\ 0 & 3\sqrt{2} & 0 & 0 \\ 0 & 0 & -4\sqrt{2} & 0 \\ 0 & 0 & 0 & -3\sqrt{2} \end{bmatrix} \frac{1}{2}\begin{bmatrix} \sqrt{2} & 0 & 1 & 1 \\ 0 & \sqrt{2} & -1 & 1 \\ -\sqrt{2} & 0 & 1 & 1 \\ 0 & -\sqrt{2} & -1 & 1 \end{bmatrix} =
$$

□

Jak spočíst SVD rozklad? Existuje celá řada přibližných metod. Zde si uvedeme (výpočtově velmi nákladný) postup pomocí Householderových transformací a Givensových rotací.

Listing 9: matrix_factorization_test.m

```
n = 10; m = 8;
% QR test
A = rand(m,n);
[R,Q] = householder(A,'qr');
disp(num2str(norm(A-Q*R)/norm(A)));
% tridiagonalization test
A = rand(n); A = (A+A')/2;
[T,P] = householder(A,'tridiag');
disp(num2str(norm(A-P*T*P')/norm(A)));
% SVD test
A = rand(m,n);
[T,P1,P2] = householder(A,'bidiag');
disp(num2str(norm(A-P1*T*P2)/norm(A)));
```

Listing 10: householder.m

```
function [A,P,P2] = householder(A,type)
m  = size(A,1); n  = size(A,2);  mn = min([m n]);
P  = eye(m);
P2 = [];
if nargin<2, type='qr'; end % type ... {'qr','tridiag','bidiag'}
switch type
  case 'qr'
    for iii=1:mn-1
      x = zeros(m,1);     normxe = zeros(m,1);
      x(iii:end,1) = A(iii:end,iii);
      normxe(iii,1) = norm(x);
      u = x - normxe;
      w = u/norm(u);
      P_ = eye(m)-2*w*(w');     P = P*P_;     A = P_*A;
    end
  case 'tridiag'
    for iii=1:mn-1
      alpha = -sign(A(iii+1,iii))*sqrt(sum(A(iii+1:end,iii).^2));
      r = sqrt( (1/2)*alpha*(alpha-A(iii+1,iii)) );     w = zeros(m,1);
      w(iii+1)     = (A(iii+1    ,iii)-alpha)/(2*r);
      w(iii+2:end) = (A(iii+2:end,iii)      )/(2*r);
      P_ = eye(m)-2*w*(w');     P = P*P_;     A = P_*A*P_;
    end
  case 'bidiag'
    P2 = eye(n);
    for iii=1:mn-1
      % row operations
      x = zeros(m,1);                         normxe = zeros(m,1);
```

```
      x(iii:end,1) = A(iii:end,iii);     normxe(iii,1) = norm(x);
      u = x - normxe;                    w = u/norm(u);
      P_ = eye(m)-2*w*(w');     P = P*P_;     A = P_*A;
      % column operations
      alpha = -sign(A(iii,iii+1))*sqrt(sum(A(iii,iii+1:end).^2));
      r = sqrt( (1/2)*alpha*(alpha-A(iii,iii+1)) );     w = zeros(n,1);
      w(iii+1)     = (A(iii,iii+1    )-alpha)/(2*r);
      w(iii+2:end) = (A(iii,iii+2:end)      )/(2*r);
      P2_= eye(n)-2*w*(w');    P2=P2_*P2;   A = A*P2_;
    end
  otherwise
    error('householder(A,opt) ... opt: ''qr'' or ''tridiag'' or ''bidiag''');
end
```

4 Kvadratické programování

4.1 Opakování z diferenciálního počtu funkcí více proměnných

Doporučujeme prostudovat [4]. Mějme $f : \mathbb{R}^n \to \mathbb{R}$. Parciální derivace funkce f podle proměnné x_i je definována jako

$$(\partial_{x_i} f)(\mathbf{a}) = \lim_{h \to 0} \frac{f(\mathbf{a} + h\mathbf{e}_i) - f(\mathbf{a})}{h}.$$

Gradient funkce f je definován jako

$$\nabla f(\mathbf{a}) = \begin{bmatrix} \partial_{x_1} f \\ \vdots \\ \partial_{x_n} f \end{bmatrix} (\mathbf{a}).$$

Příklad 13. Spočítejte gradient funkce $f : \mathbb{R}^n \to \mathbb{R}$ dané předpisem

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathsf{A}\mathbf{x} - \mathbf{x}^\top \mathbf{b},$$

kde A je symetrická matice.

Řesení. Spočtěme si nejdříve parciální derivaci podle i-té proměnné. Pro tento účel rozepíšeme f tak, aby i-tá proměnná nebyla nikde v sumě

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{2}\sum_{k=1}^{n}\sum_{l=1}^{n} x_k a_{k,l} x_l - \sum_{k=1}^{n} x_k b_k \\ &= \frac{1}{2}\sum_{\substack{k=1 \\ k\neq i}}^{n}\sum_{\substack{l=1 \\ l\neq i}}^{n} x_k a_{k,l} x_l + \frac{1}{2}\sum_{\substack{l=1 \\ l\neq i}}^{n} {\color{red}x_i} a_{i,l} x_l + \frac{1}{2}\sum_{\substack{k=1 \\ k\neq i}}^{n} x_k a_{k,i} {\color{red}x_i} + \frac{1}{2}{\color{red}x_i} a_{i,i} {\color{red}x_i} - \sum_{\substack{k=1 \\ k\neq i}}^{n} x_k b_k - {\color{red}x_i} b_i \end{aligned}$$

pak zřejmě

$$\partial_{x_i} f(\mathbf{x}) = \frac{1}{2}\sum_{\substack{l=1 \\ l\neq i}}^{n} a_{i,l} x_l + \frac{1}{2}\sum_{\substack{k=1 \\ k\neq i}}^{n} a_{k,i} x_k + a_{i,i} {\color{red}x_i} - b_i = \sum_{k=1}^{n} a_{i,k} x_k - b_i = [\mathsf{A}\mathbf{x} - \mathbf{b}]_i$$

a tedy

$$\nabla f(\mathbf{x}) = \mathsf{A}\mathbf{x} - \mathbf{b}.$$

$\square$

4.2 Metoda sdružených gradientů

Inženýrské problémy často vedou na úlohy minimalizace celkové energie systému. Energii systému zpravidla lze popsat právě vícedimenzionální kvadratickou funkcí, tzv. kvadratického funkcionálu. Často tedy hledáme řešení úlohy minimalizace kvadratického funkcionálu (matice A je pozitivně definitní). Proberme si nejprve jednodušší variantu, kdy k minimalizaci nedodáváme další omezující podmínky. Pro řešení úlohy

$$\mathbf{x}^\star = \arg\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}), \quad f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathsf{A}\mathbf{x} - \mathbf{x}^\top \mathbf{b}$$

platí (jelikož je f hladká funkce)

$$\nabla f(\mathbf{x}^\star) = \mathsf{A}\mathbf{x}^\star - \mathbf{b} = \mathbf{0} \iff \mathbf{x}^\star = \mathsf{A}^{-1}\mathbf{b}.$$

Potřebujeme tedy řešit soustavu $\mathsf{A}\mathbf{x}^\star = \mathbf{b}$, například Gaussovou eliminační metodou. Pokud je dimenze matice A velká, je provedení Gaussovy eliminace výpočetně náročné. Uvedeme si metodu, která k výpočtu $\mathbf{x}^\star$ potřebuje pouze implementovanou funkci násobení maticí A zleva. K tomu se ale dopracujeme postupnými krůčky.

Připomeňme si, že zobrazení $(\mathbf{u},\mathbf{v})_\mathsf{A} := \mathbf{u}^\top \mathsf{A}\mathbf{v}$ je skalární součin. Předpokládejme, že máme ortogonální bázi

$$(\mathbf{p}_1,\dots,\mathbf{p}_n), \quad i \neq j \;\Rightarrow\; (\mathbf{p}_i,\mathbf{p}_j)_\mathsf{A} = 0, \quad (\mathbf{p}_i,\mathbf{p}_i)_\mathsf{A} > 0$$

vzhledem k tomuto skalárnímu součinu. Když využijeme zápisu libovolného vektoru $\mathbf{x} \in \mathbb{R}^n$ v této bázi $\mathbf{x} = \sum_{i=1}^n \tilde{\alpha}_i \mathbf{p}_i$, lze minimizační úloha převést na

$$\begin{aligned}
\mathbf{x}^\star = \sum_{i=1}^n \tilde{\alpha}_i^\star \mathbf{p}_i &= \arg \min_{\sum_{i=1}^n \tilde{\alpha}_i \mathbf{p}_i \in \mathbb{R}^n} f(\sum_{i=1}^n \tilde{\alpha}_i \mathbf{p}_i) \\
&= \arg \min_{\substack{\tilde{\alpha}_i \in \mathbb{R} \\ i=1,\dots,n}} \frac{1}{2} \left(\sum_{i=1}^n \tilde{\alpha}_i \mathbf{p}_i \right)^\top \mathsf{A} \left(\sum_{j=1}^n \tilde{\alpha}_j \mathbf{p}_j \right) - \left(\sum_{i=1}^n \tilde{\alpha}_i \mathbf{p}_i \right)^\top \mathbf{b} \\
&= \arg \min_{\substack{\tilde{\alpha}_i \in \mathbb{R} \\ i=1,\dots,n}} \sum_{i=1}^n \left(\frac{1}{2} \tilde{\alpha}_i^2 \mathbf{p}_i^\top \mathsf{A}\mathbf{p}_i - \tilde{\alpha}_i \mathbf{p}_i^\top \mathbf{b} \right) = \arg \min_{\tilde{\alpha}_1 \in \mathbb{R}} f(\tilde{\alpha}_1 \mathbf{p}_1) + \dots + \min_{\tilde{\alpha}_n \in \mathbb{R}} f(\tilde{\alpha}_n \mathbf{p}_n),
\end{aligned}$$

čímž byl původní minimizační problém převeden na n nezávislých jednodimenzionálních kvadratických rovnic. Souřadnice řešení

$$\mathbf{x}^\star = \sum_{i=1}^n \tilde{\alpha}_i^\star \mathbf{p}_i, \quad 0 = \frac{df(\tilde{\alpha}\mathbf{p}_i)}{d\tilde{\alpha}}|_{\tilde{\alpha}=\tilde{\alpha}_i^\star} = \tilde{\alpha}_i^\star \mathbf{p}_i^\top \mathsf{A}\mathbf{p}_i - \mathbf{p}_i^\top \mathbf{b} \quad \Rightarrow \; \tilde{\alpha}_i^\star = \frac{\mathbf{p}_i^\top \mathbf{b}}{\mathbf{p}_i^\top \mathsf{A}\mathbf{p}_i}$$

v bázi $(\mathbf{p}_1,\dots,\mathbf{p}_n)$ můžeme spočíst paralelně, avšak pro každou souřadnici musíme spočítat $\mathsf{A}\mathbf{p}_i$, tedy násobení maticí. Pro velká n je tedy dopočtení přesného řešení $\mathbf{x}^\star$ příliš náročné. Omezme se na úlohu nalezení minima pouze na podmnožině

$$\mathbf{x}_k^\bullet \in \mathscr{S}_k = \{\mathbf{x}_0 + \tilde{\alpha}_1 \mathbf{p}_1 + \dots + \tilde{\alpha}_k \mathbf{p}_k\} = \mathbf{x}_0 + \langle \mathbf{p}_1,\dots,\mathbf{p}_k \rangle \subset \mathbb{R}^n$$

tedy

$$\begin{aligned}
\hat{\mathbf{x}}_k = \mathbf{x}_0 + \sum_{i=1}^k \tilde{\alpha}_i^\bullet \mathbf{p}_i &= \arg \min_{\sum_{i=1}^k \tilde{\alpha}_i \mathbf{p}_i \in \mathbb{R}^n} f(\mathbf{x}_0 + \sum_{i=1}^k \tilde{\alpha}_i \mathbf{p}_i) \\
&= \arg \min_{\substack{\tilde{\alpha}_i \in \mathbb{R} \\ i=1,\dots,n}} f(\mathbf{x}_0) + \sum_{i=1}^k \left(\frac{1}{2} \tilde{\alpha}_i^2 \mathbf{p}_i^\top \mathsf{A}\mathbf{p}_i + \tilde{\alpha}_i \mathbf{p}_i^\top \mathsf{A}\mathbf{x}_0 - \tilde{\alpha}_i \mathbf{p}_i^\top \mathbf{b} \right) \\
&= \arg \min_{\substack{\tilde{\alpha}_i \in \mathbb{R} \\ i=1,\dots,n}} \sum_{i=1}^k \left(\frac{1}{2} \tilde{\alpha}_i^2 \mathbf{p}_i^\top \mathsf{A}\mathbf{p}_i - \tilde{\alpha}_i \mathbf{p}_i^\top (\mathbf{b} - \mathsf{A}\mathbf{x}_0) \right) \\
&= \arg f(\mathbf{x}_0) + \min_{\tilde{\alpha}_1 \in \mathbb{R}} f_0(\tilde{\alpha}_1 \mathbf{p}_1) + \dots + \min_{\tilde{\alpha}_k \in \mathbb{R}} f_0(\tilde{\alpha}_k \mathbf{p}_k),
\end{aligned}$$

kde

$$f_0(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathsf{A}\mathbf{x} - \mathbf{x}^\top(\mathbf{b} - \mathsf{A}\mathbf{x}_0) = \frac{1}{2}\mathbf{x}^\top \mathsf{A}\mathbf{x} + \mathbf{x}^\top \mathbf{g}_0$$

a proto

$$\mathbf{x}_k^\bullet = \sum_{i=1}^k \tilde{\alpha}_i^\bullet \mathbf{p}_i, \quad 0 = \frac{df_0(\tilde{\alpha}\mathbf{p}_i)}{d\tilde{\alpha}}|_{\tilde{\alpha}=\tilde{\alpha}_i^\bullet} = \tilde{\alpha}_i^\bullet \mathbf{p}_i^\top \mathsf{A}\mathbf{p}_i - \mathbf{p}_i^\top \mathbf{g}_0 \quad \Rightarrow \; \tilde{\alpha}_i^\bullet = -\frac{\mathbf{p}_i^\top \mathbf{g}_0}{\mathbf{p}_i^\top \mathsf{A}\mathbf{p}_i}.$$

Všimněme si, že minimum na $\mathscr{S}_k$ můžeme získat z minima na $\mathscr{S}_{k-1}$ pouze dořešením jednodimenzionální úlohy, protože

$$f(\mathbf{x}_k^\bullet) = \min_{\mathbf{x}\in\mathscr{S}_k} f(\mathbf{x}) = f(\mathbf{x}_{k-1}^\bullet) + \min_{\tilde{\alpha}_k\in\mathbb{R}} f_0(\tilde{\alpha}_k\mathbf{p}_k)$$

a můžeme se tedy pokusit generovat generovat posloupnost $\mathbf{x}_k^\bullet$, $k = 1,\dots$ iterativně. Protože chceme, aby zde uvedené vzorce korespondovaly s literaturou, přejdeme od koeficientů $\tilde{\alpha}$ k $\alpha = -\tilde{\alpha}$

$$\mathbf{x}_k^\bullet = \mathbf{x}_{k-1}^\bullet - \alpha_k\mathbf{p}_k, \quad \alpha_k = \frac{\mathbf{g}_0^\top\mathbf{p}_k}{\mathbf{p}_k^\top\mathsf{A}\mathbf{p}_k} = \frac{\mathbf{g}_{k-1}^\top\mathbf{p}_k}{\mathbf{p}_k^\top\mathsf{A}\mathbf{p}_k}$$

jelikož lze upravit

$$\begin{aligned}\mathbf{g}_{k-1}^\top\mathbf{p}_k &= \left(\mathsf{A}\mathbf{x}_{k-1}^\bullet - \mathbf{b}\right)^\top\mathbf{p}_k = \left(\mathsf{A}\left(\mathbf{x}_{k-2}^\bullet - \alpha_{k-1}\mathbf{p}_{k-1}\right) - \mathbf{b}\right)^\top\mathbf{p}_k = \left(\mathsf{A}\mathbf{x}_{k-2}^\bullet - \mathbf{b}\right)^\top\mathbf{p}_k - \alpha_{k-1}\overbrace{\mathbf{p}_{k-1}^\top\mathsf{A}\mathbf{p}_k}^{=0}\\ &= \mathbf{g}_{k-2}^\top\mathbf{p}_k = \dots = \mathbf{g}_0^\top\mathbf{p}_k.\end{aligned}$$

Zatím jsme si nic neříkali o vektorech $\mathbf{p}_i$, které bychom také chtěli generovat iterativně. Máme vektory $\mathbf{p}_1,\dots,\mathbf{p}_k$ a chtěli bychom vygenerovat nový $\mathbf{p}_{k+1}$. Vzpomeneme-li si na Garmmův–Schmidtův ortogonalizační proces (viz Příklad 7), tak pokud máme

$$\mathbf{h}_k \notin \langle\mathbf{p}_1,\dots,\mathbf{p}_k\rangle, \quad \text{pak } \mathbf{p}_{k+1} = \mathbf{h}_k + \beta_{k1}\mathbf{p}_1 + \dots + \beta_{kk}\mathbf{p}_k, \quad \text{kde } \beta_{ki} = -\frac{\mathbf{p}_i^\top\mathsf{A}\mathbf{h}_k}{\mathbf{p}_i^\top\mathsf{A}\mathbf{p}_i},$$

protože

$$0 = \mathbf{p}_i^\top\mathsf{A}\mathbf{p}_{k+1} = \mathbf{p}_i^\top\mathsf{A}\mathbf{h}_k + \beta_{k1}\mathbf{p}_i^\top\mathsf{A}\mathbf{p}_1 + \dots + \beta_{kk}\mathbf{p}_i^\top\mathsf{A}\mathbf{p}_k = \mathbf{p}_i^\top\mathsf{A}\mathbf{h}_k + \beta_{ki}\mathbf{p}_i^\top\mathsf{A}\mathbf{p}_i.$$

Takto bychom uchovávali všechny vektory $\mathbf{p}_1,\dots,\mathbf{p}_k$, a potřebovali vypočíst k skalárních součinů. Pokud ale budeme postupovat chytře, lze obě tyto nevýhody eliminovat. Hledejme

$$\mathbf{x}_k^\bullet \in \mathbf{x}_0 + \mathscr{K}_k, \quad \mathscr{K}_k = \langle\mathbf{g}_0, \mathsf{A}\mathbf{g}_0,\dots,\mathsf{A}^{k-1}\mathbf{g}_0\rangle.$$

Ukažme si, že gradient $\mathbf{g}_k = \mathbf{g}(\mathbf{x}_k^\bullet)$ v $\mathbf{x}_k^\bullet$ je kolmý k prostoru $\mathscr{K}_k$, tedy

$$\mathbf{p}^\top\mathbf{g}_k = 0 \quad \forall\mathbf{p}\in\mathscr{K}_k.$$

Jelikož jsme $\mathbf{x}_k^\bullet$ zavedli jako $f(\mathbf{x}_k^\bullet) = \min_{\mathbf{x}_0+\mathscr{K}_k} f(\mathbf{x})$ tak pro všechna $\mathbf{d}\in\mathscr{K}_k$ platí

$$\begin{aligned}\frac{1}{2}\left(\mathbf{x}_k^\bullet\right)^\top\mathsf{A}\mathbf{x}_k^\bullet - \left(\mathbf{x}_k^\bullet\right)^\top\mathbf{b} = f(\mathbf{x}_k^\bullet) \le f(\mathbf{x}_k^\bullet + \varepsilon\mathbf{d}) &= \frac{1}{2}\left(\mathbf{x}_k^\bullet + \varepsilon\mathbf{d}\right)^\top\mathsf{A}\left(\mathbf{x}_k^\bullet + \varepsilon\mathbf{d}\right) - \left(\mathbf{x}_k^\bullet + \varepsilon\mathbf{d}\right)^\top\mathbf{b}\\ &= \frac{1}{2}\left(\mathbf{x}_k^\bullet\right)^\top\mathsf{A}\mathbf{x}_k^\bullet + \frac{\varepsilon^2}{2}\mathbf{d}^\top\mathsf{A}\mathbf{d} + \varepsilon\mathbf{d}^\top\mathsf{A}\mathbf{x}_k^\bullet - \left(\mathbf{x}_k^\bullet\right)^\top\mathbf{b} - \varepsilon\mathbf{d}^\top\mathbf{b}\\ &= f(\mathbf{x}_k^\bullet) + \frac{\varepsilon^2}{2}\mathbf{d}^\top\mathsf{A}\mathbf{d} + \varepsilon\mathbf{d}^\top\underbrace{\left(\mathsf{A}\mathbf{x}_k^\bullet - \mathbf{b}\right)}_{\mathbf{g}_k}\end{aligned}$$

Podívejme se na člen $\varepsilon\left(\frac{\varepsilon}{2}\mathbf{d}^\top\mathsf{A}\mathbf{d} + \mathbf{d}^\top\mathbf{g}_k\right)$, který nesmí být záporný. Pro pevné $\mathbf{d}\in\mathscr{K}_k$ jsou $\mathbf{d}^\top\mathsf{A}\mathbf{d} > 0$ a $\mathbf{d}^\top\mathbf{g}_k$ reálná čísla. Pokud je číslo $\mathbf{d}^\top\mathbf{g}_k$ nenulové, pak volbou

$$\varepsilon = -\frac{\mathbf{d}^\top\mathbf{g}_k}{\mathbf{d}^\top\mathsf{A}\mathbf{d}} \text{ dostáváme } \varepsilon\left(\frac{\varepsilon}{2}\mathbf{d}^\top\mathsf{A}\mathbf{d}+\right) = -\frac{\left(\mathbf{d}^\top\mathbf{g}_k\right)^2}{2\mathbf{d}^\top\mathsf{A}\mathbf{d}} < 0,$$

což nelze (v $\mathbf{x}_k^\bullet$ je minimum). Nutně tedy $\mathbf{d}^\top\mathbf{g}_k = 0$ pro všechny $\mathbf{d}\in\mathscr{K}_k$.

Pokud $\mathbf{g}_k \neq \mathbf{0}$, pak $\mathbf{g}_k \notin \mathscr{K}_k$ a jelikož $\mathbf{g}_k \in \mathscr{K}_{k+1}$ je ideálním kandidátem pro $\mathbf{h}_k$. Navíc je $\mathbf{g}_k$ A–kolmý k $\mathscr{K}_{k-1}$

$$\mathbf{p} \in \mathscr{K}_{k-1} \;\Rightarrow\; \mathbf{q} = \mathsf{A}\mathbf{p} \in \mathscr{K}_k \;\Rightarrow\; 0 = \mathbf{q}^\top \mathbf{g}_k = \mathbf{p}^\top \mathsf{A}\mathbf{g}_k \;\Rightarrow\; \mathbf{g}_k \perp_{\mathsf{A}} \mathscr{K}_k.$$

tedy

$$\mathbf{p}_{k+1} = \mathbf{g}_k + \beta_k \mathbf{p}_k, \quad \beta_k = -\frac{\mathbf{p}_k^\top \mathsf{A}\mathbf{g}_k}{\mathbf{p}_k^\top \mathsf{A}\mathbf{p}_k}.$$

Nyní ještě uvedeme několik zjednodušujících přepisů

$$\mathbf{g}_{k-1}^\top \mathbf{p}_k = \mathbf{g}_{k-1}^\top(\mathbf{g}_{k-1} + \beta_{k-1}\mathbf{p}_{k-1}) = \|\mathbf{g}_{k-1}\|^2 + \beta_{k-1} \underbrace{\mathbf{g}_{k-1}^\top \mathbf{p}_{k-1}}_{=0} = \|\mathbf{g}_{k-1}\|^2$$

$$\mathbf{x}_k^\bullet = \mathbf{x}_{k-1}^\bullet - \alpha_k \mathbf{p}_k \;\Rightarrow\; \mathsf{A}\mathbf{x}_k^\bullet - \mathbf{b} = \mathsf{A}\mathbf{x}_{k-1}^\bullet - \mathbf{b} - \alpha_k \mathsf{A}\mathbf{p}_k \;\Rightarrow\; \mathsf{A}\mathbf{p}_k = \frac{1}{\alpha_k}(\mathbf{g}_{k-1} - \mathbf{g}_k)$$

$$\alpha_k \mathbf{g}_k^\top \mathsf{A}\mathbf{p}_k = \mathbf{g}_k^\top(\mathbf{g}_{k-1} - \mathbf{g}_k) = -\|\mathbf{g}_k\|^2, \quad \text{protože } \mathbf{g}_{k-1} \in \mathscr{K}_k,\ \mathbf{g}_k \perp \mathscr{K}_k$$

$$\alpha_k \mathbf{p}_k^\top \mathsf{A}\mathbf{p}_k = \alpha_k \mathbf{p}_k^\top \frac{1}{\alpha_k}(\mathbf{g}_{k-1} - \mathbf{g}_k) = \mathbf{p}_k^\top \mathbf{g}_{k-1} - \mathbf{p}_k^\top \mathbf{g}_k = \|\mathbf{g}_{k-1}\|^2, \quad \text{protože } \mathbf{p}_k \in \mathscr{K}_k.$$

Nyní již můžeme zapsat algoritmus metody sdružených gradientů.

- Z počáteční inicializace $\mathbf{x}_0$ spočteme $\mathbf{g}_0$ a položíme $\mathbf{p}_0 = \mathbf{g}_0$.
- Jestliže známe $\mathbf{x}_{k-1}$ a $\mathbf{g}_{k-1}$ spočteme

$$\begin{aligned}
\mathbf{x}_k &= \mathbf{x}_{k-1} - \alpha_k \mathbf{p}_k, \quad \alpha_k = \frac{\mathbf{g}_{k-1}^\top \mathbf{p}_k}{\mathbf{p}_k^\top \mathsf{A}\mathbf{p}_k} = \frac{\|\mathbf{g}_{k-1}\|^2}{\mathbf{p}_k^\top \mathsf{A}\mathbf{p}_k} \\
\mathbf{g}_k &= \mathbf{g}_{k-1} - \alpha_k \mathsf{A}\mathbf{p}_k \\
\mathbf{p}_{k+1} &= \mathbf{g}_k + \beta_k \mathbf{p}_k, \quad \beta_k = -\frac{\mathbf{p}_k^\top \mathsf{A}\mathbf{g}_k}{\mathbf{p}_k^\top \mathsf{A}\mathbf{p}_k} = \frac{\|\mathbf{g}_k\|^2}{\|\mathbf{g}_{k-1}\|^2}.
\end{aligned}$$

Zdrojový kód 11 metody sdružených gradientů otestujeme na 2×2 příkladu, viz 12.

Listing 11: cg_method.m

```
function x = cg_method( A, b, x0, tol)
% x = cg_method( A, b, x0, tol)
% initialization
x = x0;   g = A*x-b;   p =g;   old_g_norm = norm(g);
while old_g_norm > tol
  Ap = A*p;
  alpha = old_g_norm^2/(p'*Ap);
  x = x - alpha*p;
  g = g - alpha*Ap;
  beta = (norm(g)^2)/(old_g_norm^2);
  p = g + beta*p;
  old_g_norm = norm(g);
end
```

Listing 12: cg_test.m

```
A = [5 2;2 5]; b = [-5 19]';
x_dir = A\b;
x_cg  = cg_method(A,b,0*b,1e-16);
```

4.3 Dualita v QP

Mějme úlohu

$$\min_{\mathbf{x}\in\Omega} f(\mathbf{x}), \quad f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathsf{A}\mathbf{x} - \mathbf{b}^\top\mathbf{x}, \quad \Omega = \{\mathbf{x}|\ \mathsf{B}_I\mathbf{x} \le \mathbf{c}_I\ \wedge\ \mathsf{B}_E\mathbf{x} = \mathbf{c}_E\}$$

Řešení $\overline{\mathbf{x}}$ této úlohy musí splňovat

$$\begin{aligned}
\mathsf{A}\overline{\mathbf{x}} - \mathbf{b} + \mathsf{B}^\top\lambda &= \mathbf{0}\\
\overline{\mathbf{x}} \in \Omega,\ \text{tedy}\ \mathsf{B}_I\overline{\mathbf{x}} - \mathbf{c}_I \le 0\ &\wedge\ \mathsf{B}_E\overline{\mathbf{x}} - \mathbf{c}_E = 0\\
\lambda = \begin{bmatrix}\lambda_I\\ \lambda_E\end{bmatrix}, \quad \lambda_I \ge 0, &\quad \mathsf{B} = \begin{bmatrix}\mathsf{B}_I\\ \mathsf{B}_E\end{bmatrix}\\
\lambda^\top\mathsf{B}\left(\overline{\mathbf{x}} - \mathbf{c}\right) &= 0
\end{aligned}$$

Gradient f je

$$\mathbf{g}(\mathbf{x}) = \nabla f(\mathbf{x}) = \mathsf{A}\mathbf{x} - \mathbf{b}$$

Lagrangián

$$\begin{aligned}
\mathscr{L}(\mathbf{x},\lambda) &= f(\mathbf{x}) + \lambda^\top(\mathsf{B}\mathbf{x} - \mathbf{c})\\
\nabla_{\mathbf{x}}\mathscr{L}(\mathbf{x},\lambda) &= \nabla f(\mathbf{x}) + \mathsf{B}^\top\lambda = \mathsf{A}\mathbf{x} - \mathbf{b} + \mathsf{B}^\top\lambda\\
\nabla_{\lambda}\mathscr{L}(\mathbf{x},\lambda) &= \mathsf{B}\mathbf{x} - \mathbf{c}
\end{aligned}$$

Sedlobodová úloha: najdi $(\overline{\mathbf{x}},\overline{\lambda}) \in \mathbb{R}^n \times \Psi$

$$\mathscr{L}(\overline{\mathbf{x}},\lambda) \le \mathscr{L}(\overline{\mathbf{x}},\overline{\lambda}) \le \mathscr{L}(\mathbf{x},\overline{\lambda})$$

Podmínky sedlového bodu

$$\begin{aligned}
&\nabla_{\mathbf{x}}\mathscr{L}(\overline{\mathbf{x}},\overline{\lambda}) = \mathbf{0}\\
&\left[\nabla_{\lambda}\mathscr{L}(\overline{\mathbf{x}},\overline{\lambda})\right]_E = \mathbf{0},\ \left[\nabla_{\lambda}\mathscr{L}(\overline{\mathbf{x}},\overline{\lambda})\right]_I \le \mathbf{0}\\
&\overline{\lambda}^\top\nabla_{\lambda}\mathscr{L}(\overline{\mathbf{x}},\overline{\lambda}) = 0
\end{aligned}$$

4.3.1 Dualita pro rovnostní omezení

Mějme úlohu

$$\min_{\mathbf{x}\in\Omega} f(\mathbf{x}), \quad f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top\mathsf{A}\mathbf{x} - \mathbf{b}^\top\mathbf{x}, \quad \Omega_E = \{\mathbf{x}|\ \mathsf{B}\mathbf{x} = \mathbf{c}\}$$

tedy sedlobodovou úlohu najdi $(\overline{\mathbf{x}},\overline{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^m$

$$\mathscr{L}(\overline{\mathbf{x}},\lambda) \le \mathscr{L}(\overline{\mathbf{x}},\overline{\lambda}) \le \mathscr{L}(\mathbf{x},\overline{\lambda}), \quad \text{kde}\ \mathscr{L}(\mathbf{x},\lambda) = \frac{1}{2}\mathbf{x}^\top\mathsf{A}\mathbf{x} - \mathbf{b}^\top\mathbf{x} + \lambda^\top(\mathsf{B}\mathbf{x} - \mathbf{c})$$

Pro fixní $\hat{\lambda}$ dopočteme $\mathbf{x}(\hat{\lambda}) = \mathsf{A}^{-1}\left(\mathbf{b} - \mathsf{B}^\top\hat{\lambda}\right)$.

Označme si funkci

$$\begin{aligned}
\Theta(\lambda) &= \min_{\mathbf{x}\in\mathbb{R}^n}\mathscr{L}(\mathbf{x},\lambda) = \mathscr{L}(\mathbf{x}(\lambda),\lambda) =\\
&= \frac{1}{2}\left(\mathbf{b} - \mathsf{B}^\top\lambda\right)^\top\mathsf{A}^{-1}\mathsf{A}\mathsf{A}^{-1}\left(\mathbf{b} - \mathsf{B}^\top\lambda\right) - \mathbf{b}^\top\mathsf{A}^{-1}\left(\mathbf{b} - \mathsf{B}^\top\lambda\right) + \lambda^\top\left(\mathsf{B}\mathsf{A}^{-1}\left(\mathbf{b} - \mathsf{B}^\top\lambda\right) - \mathbf{c}\right)\\
&= \frac{1}{2}\lambda^\top\mathsf{B}\mathsf{A}^{-1}\mathsf{B}^\top\lambda + \lambda^\top\left(\mathsf{B}\mathsf{A}^{-1}\mathbf{b} - \mathbf{c}\right) - \frac{1}{2}\mathbf{b}^\top\mathsf{A}^{-1}\mathbf{b}
\end{aligned}$$

tuto funkci minimalizujeme na $\mathbb{R}^m$, tedy

$$\mathbf{0} = \nabla\Theta(\overline{\lambda}) = \mathsf{B}\mathsf{A}^{-1}\mathsf{B}^\top\overline{\lambda} - \left(\mathbf{c} - \mathsf{B}\mathsf{A}^{-1}\mathbf{b}\right)$$

4.3.2 Dualita pro nerovnostní omezení

dsfsd

4.4 Rozšířený Lagrangián

Vezměme si problém

$$\mathbf{x}^\star = \arg\min_{\mathsf{B}\mathbf{x}=\mathbf{c}} \frac{1}{2}\mathbf{x}^\top \mathsf{A}\mathbf{x} - \mathbf{x}^\top \mathbf{b}$$

Reference

[1] Z. Dostál, V. Vondrák: Lineární algebra. VŠB (2012),
http://mi21.vsb.cz/modul/linearni-algebra

[2] Z. Dostál, L. Šindel: Lineární algebra pro kombinované a distanční studium. VŠB (2003),
http://homel.vsb.cz/~s1a64/lait/linearni_algebra_pro_kombinovane_studium.pdf

[3] T. Kozubek, T. Brzobohatý, M. Jarošová, V. Hapla, A. Markopoulos: Lineární Algebra s MatLabem. VŠB (2012)
http://mi21.vsb.cz/sites/mi21.vsb.cz/files/unit/linearni_algebra_s_matlabem.pdf

[4] J. Kuben, Š. Mayerová, P. Račková, P. Šarmanová: Diferenciální počet funkcí více proměnných. VŠB (2012),
http://mi21.vsb.cz/modul/integralni-pocet-funkci-vice-promennych