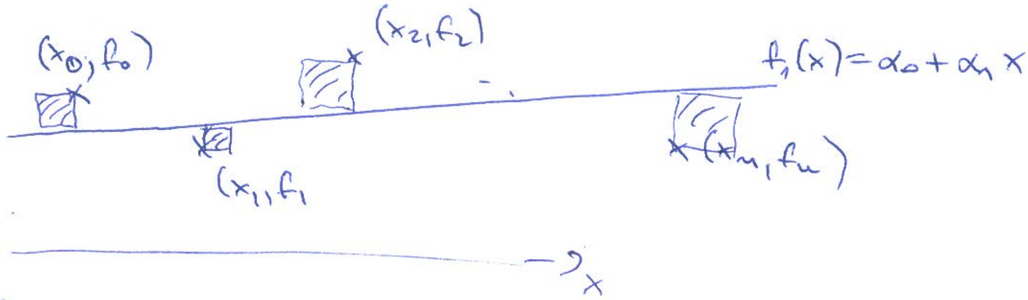
NM – Aproximace metodou nejmenších čtverců

Lineární regrese v ℓ^2-normě

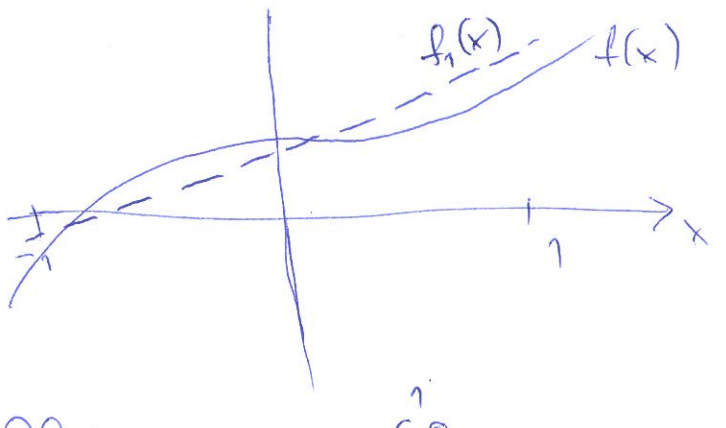


$$\begin{cases} \text{Hledáme } f_1(x)=\alpha_0+\alpha_1 x \in P_1 : \\ \underbrace{\sum_{j=0}^{m}\left[f_j-\alpha_0-\alpha_1 x\right]^2}_{=:Q(\alpha_0,\alpha_1)} \longrightarrow \min_{\alpha_0,\alpha_1\in\mathbb{R}}! \end{cases}$$

$=: Q(\alpha_0,\alpha_1)$ .. je ryze konvexní kvadratická funkce,
tj. úloha má právě jedno řešení $\bar\alpha : \nabla Q(\bar\alpha)=\bar 0$

$$\left.\begin{aligned} \frac{\partial Q}{\partial \alpha_0}(\alpha_0,\alpha_1) &= \sum_{j=0}^{m} 2\cdot\left[f_j-\alpha_0-\alpha_1 x_j\right]\cdot(-1) \overset{!}{=} 0 \\ \frac{\partial Q}{\partial \alpha_1}(\alpha_0,\alpha_1) &= \sum_{j=0}^{m} 2\cdot\left[f_j-\alpha_0-\alpha_1 x_j\right]\cdot(-x_j) \overset{!}{=} 0 \end{aligned}\right\} \quad \boxed{\begin{bmatrix} \sum_{j=0}^m 1 , & \sum_{j=0}^m x_j \\ \sum_{j=0}^m x_j , & \sum_{j=0}^m (x_j)^2 \end{bmatrix}\begin{bmatrix}\alpha_0\\ \alpha_1\end{bmatrix} = \begin{bmatrix}\sum_{j=0}^m f_j \\ \sum_{j=0}^m x_j f_j\end{bmatrix}}$$

Lineární regrese v L^2-normě



$$\begin{cases} \text{Hledáme } f_1(x)=\alpha_0+\alpha_1 x \in P_1 : \\ \underbrace{\int_{-1}^{1}\left[f(x)-\alpha_0-\alpha_1 x\right]^2 dx}_{} \longrightarrow \min_{\alpha_0,\alpha_1\in\mathbb{R}}! \end{cases}$$

$= \boxed{Q(\alpha_0,\alpha_1) \text{ je opět ryze konvex. kvadrat. funkce. Cvič.}}$

$$\left.\begin{aligned} \frac{\partial Q}{\partial \alpha_0}(\alpha_0,\alpha_1) &= -2\cdot\int_{-1}^{1}\left[f(x)-\alpha_0-\alpha_1 x\right]dx \overset{!}{=} 0 \\ & \qquad\qquad\qquad \int_{-1}^{1} x = 0 \\ \frac{\partial Q}{\partial \alpha_1}(\alpha_0,\alpha_1) &= -2\int_{-1}^{1} x\left[f(x)-\alpha_0-\alpha_1 x\right]dx \overset{!}{=} 0 \end{aligned}\right\} \quad \boxed{\begin{bmatrix} \int_{-1}^1 1 , & 0 \\ 0 , & \int_{-1}^1 x^2 \end{bmatrix}\begin{bmatrix}\alpha_0\\ \alpha_1\end{bmatrix} = \begin{bmatrix}\int_{-1}^1 f(x) \\ \int_{-1}^1 x f(x)\end{bmatrix}}$$

$1 \perp x$ ve skal. součinu $(f(x),g(x))_{L^2(-1,1)} := \int_{-1}^{1} f(x)g(x)\,dx$, nemusíme proto řešit soustavu lin. rovnic (matice je diagonální).

4.10.'16

(1)

Ortogonální systémy a metoda nejm. čtverců

Mějme vektorový prostor V nad tělesem $\mathbb{R}$ (nad $\mathbb{C}$).

Zaveďme skalární součin (u,v)

- bilineární forma

  tj. $\forall \alpha,\beta \in \mathbb{R}$ $(\in \mathbb{C})$ $\forall u,v,w \in V$:

  1) $(\alpha u+\beta v, w) = \alpha(u,w)+\beta(v,w)$

  2) $(u, \alpha v+\beta w) = \alpha(u,v)+\beta(u,w)$ v příp. $\mathbb{R}$

  $\left(= \alpha^*(u,v)+\beta^*(u,w)\right.$ v příp. $\mathbb{C}$, kde $\left.\alpha=a+ib,\ \alpha^*=a-ib\right)$

- která je pozitivně definitní

  tj. $\forall v \in V,\ v \neq 0$: $(v,v) > 0$

- a symetrická (Hermitovská v příp. $\mathbb{C}$)

  tj. $\forall u,v \in V$: $(u,v) = (v,u)$

  $\left(= (v,u)^*\right)$

Systém nenulových prvků $(\varphi_i)_{i=0}^n \subseteq V$ se nazývá ortogonální, pokud $(\varphi_i,\varphi_j)=0$ pro $i \neq j$.

Skalární součin indukuje normu $\|u\| := \sqrt{(u,u)}$.

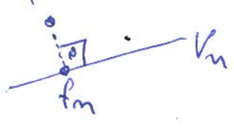
~~Metodu nejm. čtverců zo~~

$\boxed{\text{Cvičení. Ortogonální systém } (\varphi_i)_{i=0}^n \subset V \text{ tvoří bázi } \langle \varphi_0,\dots,\varphi_n\rangle.}$

Metodu nejm. čtverců zobecníme takto:

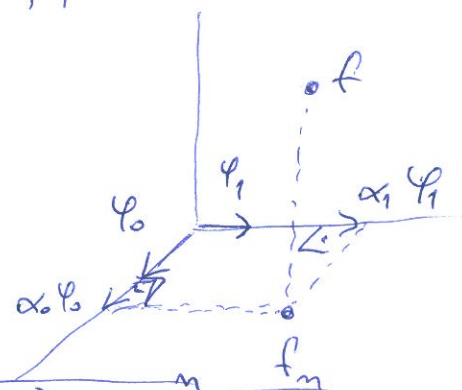
Dáno $f \in V$, $(\varphi_i)_{i=0}^n$ lin. nezávislé.

$$\left\{\begin{array}{l} \text{Hledáme } f_n = \sum_{j=0}^n \alpha_j \varphi_j : \\ \|f-f_n\|^2 \longrightarrow \min\limits_{\alpha_0,\dots,\alpha_n \in \mathbb{R}\ (\in\mathbb{C})} ! \end{array}\right. \overset{\boxed{\text{Cvič.}}}{\Longleftrightarrow} (f-f_n,\varphi_i)=0 \ \ \forall i \overset{\boxed{\text{Cvič.}}}{\Longleftrightarrow} \begin{bmatrix} (\varphi_0,\varphi_0) & \cdots & (\varphi_0,\varphi_n) \\ (\varphi_1,\varphi_0) & \cdots & (\varphi_1,\varphi_n) \\ \vdots & & \vdots \\ (\varphi_n,\varphi_0) & \cdots & (\varphi_n,\varphi_n) \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} (\varphi_0,f) \\ \vdots \\ (\varphi_n,f) \end{bmatrix}$$



Je-li báze $(\varphi_i)_{i=0}^n$ ortogonální, pak

$$\boxed{f_n = \sum_{j=0}^{n} \underbrace{\frac{(\varphi_j, f)}{(\varphi_j, \varphi_j)}}_{=\alpha_j} \cdot \varphi_j}$$



Př. Eukleidovský prostor $\mathbb{R}^n$, $\boxed{(\bar{u}, \bar{v}) := \sum_{i=1}^{n} u_i v_i}$

$(\bar{e}_j)_{j=1}^{n}$, kde $\bar{e}_j = (0, \dots, 0, \overset{j}{1}, 0, \dots, 0)$, tvoří ortog. systém

Př. Prostor polynomů $P_n := \{a_0 + a_1 x + \dots + a_n x^n \mid a_i \in \mathbb{R}\}$

$a \leq x_0 < x_1 < x_2 < \dots < x_n \leq b$

$w_0, w_1, w_2, \dots, w_n > 0$

Diskrétní ℓ^2-skalární součin $\boxed{(f, g)_{\ell^2, w} := \sum_{i=0}^{n} w_i f(x_i) g(x_i)}$

$(L_i^n(x))_{i=0}^{n}$, kde $L_i^n(x) := \dfrac{\prod_{\substack{j=0\\ j\neq i}}^{n} (x - x_j)}{\prod_{\substack{j=0\\ j\neq i}}^{n} (x_i - x_j)}$, Lagr. báze tvoří ortog. systém.

Př. Prostor spojitých funkcí $C(\langle a, b \rangle)$

$w \in C((a,b))$: $w(x) > 0$ v (a,b)

$\int_a^b w(x)\, dx \in \mathbb{R}$

L^2-skalární součin $\boxed{(f, g)_{L^2, w} := \int_a^b w(x) f(x) g(x)\, dx}$

Ortogonální bázi lze sestrojit Gram-Schmidtovým algor.

$$\boxed{\varphi_{k+1} := \underset{V_{k+1} \setminus V_k}{\underset{\cap}{v_{k+1}}} - \sum_{j=0}^{k} \beta_j^{k+1} \varphi_j, \quad \beta_j := \frac{(v_{k+1}, \varphi_j)}{(\varphi_j, \varphi_j)}}$$

Ortogonální systémy polynomů splňují 3-člennou rekurenci:

$$\boxed{p_{k+1}(x) = x\, p_k(x) - \frac{(x p_k(x), p_k(x))}{(p_k(x), p_k(x))} p_k(x) - \frac{(x p_k, p_{k-1})}{(p_{k-1}, p_{k-1})} p_{k-1}(x)} - \underbrace{\frac{(p_k, \overset{\in P_{k-1}}{x p_{k-2}})}{(p_{k-2}, p_{k-2})}}_{=0} p_{k-2} - 0 \dots$$

<u>Cvičení</u>: $a=-b\neq 0$, $w(x)$ nechť je sudá,

$$(f,g)_{L^2,w} := \int_a^b w(x)\, f(x)\, g(x)\, dx .$$

Dokažte, že v ortog. systému polynomů $p_0(x) := 1$, $p_1(x) := x$, $\vdots$

budou $p_{2j}(x)$ sudé funkce

$p_{2j+1}(x)$ liché funkce

a tedy $(x\, p_k(x), p_k(x))_{L^2,w} = \int_a^b \underbrace{w(x)}_{\text{sudá}}\, \underbrace{x\, p_k(x)\, p_k(x)}_{\text{lichá}}\, dx = 0$,

tj.
$$\boxed{p_{k+1}(x) = x\, p_k(x) - \frac{(x\, p_k, p_{k-1})}{(p_{k-1}, p_{k-1})}\, p_{k-1}(x)}$$

Př. <u>Čebyševovy polynomy</u> $T_0(x) := 1$, $T_1(x) := x$, $\boxed{T_{k+1}(x) := 2x\, T_k(x) - T_{k-1}(x)}$

jsou ortogonální vzhledem k
$$\boxed{(f,g)_{L^2,w} := \int_{-1}^{1} \underbrace{\frac{1}{\sqrt{1-x^2}}}_{w(x)}\, f(x)\, g(x)\, dx}$$

Tvoří také ortog. systém prostoru P_n vzhledem k diskrétní ℓ^2-normě v uzlech $T_{n+1}(x)$.

Př. <u>Legendrovy polynomy</u> $P_0(x) := 1$, $P_1(x) := x$, $\boxed{P_{k+1}(x) := \frac{2k+1}{k+1}\, x\, P_k(x) - \frac{k}{k+1}\, P_{k-1}(x)}$

jsou ortogonální vzhledem k
$$\boxed{(f,g)_{L^2,w=1} := \int_{-1}^{1} \underset{w}{1} \cdot f(x)\, g(x)\, dx} .$$

Př. <u>Fourierova báze</u> $\boxed{\varphi_k(x) := e^{ikx}}$, $k = 0, \dots, n$

tvoří ortogonální systém vzhledem k
$$\boxed{(f,g) := \int_0^{2\pi} f(x)\, g(x)^*\, dx},$$

Analogicky: <u>trigonometrická báze</u> $1, \cos x, \sin x, \cos(2x), \sin(2x), \dots$

tvoří ortog. systém vzhledem k $(f,g) := \int_0^{2\pi} f(x)\, g(x)\, dx$

Fourierova aproximace (transformace, pokud $n=\infty$)

$$\boxed{f_n(x) = \frac{1}{2\pi}\sum_{j=0}^{n}\Big[\underbrace{\int_0^{2\pi} f(x)\, e^{-i[jx]}\,dx}_{=\alpha_j}\Big]\, e^{i[jx]}}$$

Diskrétní Fourierova transformace (DFT)

Dáno $a_0, a_1, \dots, a_n \in \mathbb{C}$

$$\boxed{A_k := \sum_{j=0}^{n} a_j\, e^{-2\pi i\left[k\frac{j}{n+1}\right]},\ k=0,\dots,n} \quad \dots \underline{\text{složitost } O(n^2)}$$

Rychlá DFT =: FFT (jpeg, mp3, mpeg, ...)

$n =: n_L$, $\boxed{n_L + 1 = 2^L}$, $L \in \mathbb{N}$

$$A_k = \sum_{\substack{j=0 \\ j\in\{0,2,4,\dots\} \\ \text{sudé}}}^{n_L} a_j\, e^{-2\pi i\left[k\frac{j}{n_L+1}\right]} + \sum_{\substack{j=0 \\ j\in\{1,3,5,\dots\} \\ \text{liché}}}^{n_L} a_j\, e^{-2\pi i\left[k\frac{j}{n_L+1}\right]}$$

$$\boxed{n_{L-1} := \frac{n_L - 1}{2} = 2^{L-1} - 1}$$

$\boxed{\text{Cvič.}}$

$$\sum_{j=0}^{n_{L-1}} a_{2j}\, e^{-2\pi i\left[k\frac{j}{n_{L-1}+1}\right]} + e^{-2\pi i\left[k\frac{1}{n_L+1}\right]} \cdot \sum_{j=0}^{n_{L-1}} a_{2j+1}\, e^{-2\pi i\left[k\frac{j}{n_{L-1}+1}\right]}$$

tj.

$$\boxed{\begin{aligned} DFT_{L,k}\left((a_j)_{j=0}^{n_L}\right) &= DFT_{L-1,k}\left((a_{2j})_{j=0}^{n_{L-1}}\right) \\ &\quad + e^{-2\pi i\left[\frac{k}{n_L+1}\right]} \cdot DFT_{L-1,k}\left((a_{2j+1})_{j=0}^{n_{L-1}}\right) \end{aligned}}$$

$\dots$ $\underline{\text{složitost } O(n \cdot \log_2 n)} = O(L \cdot n)$