Metoda nejmenších čtverců

Pomocí skalárního součinu funkcí spojitých na intervalu $\mathrm{C}\langle a,b\rangle$ odvodíme optimální aproximaci funkce f pomocí zadaných, tzv. aproximačních funkcí $\varphi_1,\ldots,\varphi_k$, které jsou zpravidla jednoduchými, a tudíž pro další použití ,,rozumnými" funkcemi. Jedná se nejčastěji o polynomy nebo trigonometrické funkce. V případě $(b-a)$-periodického signálu v čase (může se jednat o elektromagnetický nabo akustický signál) lze aproximaci trigonometrickými funkcemi tvaru $\cos\frac{2\pi kt}{b-a}$ a $\sin\frac{2\pi kt}{b-a}$ pro $k=0,\ldots,n$ chápat jako rozklad zadaného signálu na signál s perodou $b-a$ a jeho k-té harmonické, přičemž získaná lineární kombinace k-tých harmonických pomocí MNČ vystihuje nejlépe průběh zadaného, tedy skutečného signálu.

Pro naše úvahy a odvození využijeme geometrický model. Jeho hlavní idea spočívá v tom, že budeme uvažovat vektorový prostor všech spojitých funkcí $\mathrm{C}\langle a,b\rangle$ na intervalu $\langle a,b\rangle$ (poznamenejme, že se jedná nekonečně-dimensionální prostor) a lineární obal (tzn. množinu všech lineárních konbinací) zadaných aproximačních funkcí $\varphi_1,\ldots,\varphi_k\in\mathrm{C}\langle a,b\rangle$, což je ve výchozím prostoru $\mathrm{C}\langle a,b\rangle$ vektorový podprostor konečné dimense k. Na $\mathrm{C}\langle a,b\rangle$ zavedeme jistým skalární součin předpisem

$$f*g=\int_a^b f(x)g(x)dx,$$

který nám mimo jiné umožní definovat ,,velikost" f jako $||f||=\sqrt{f*f}$ a,,vzdálenost" funkcí f,g jako $||f-g||$. Optimální aproximace f^* ve smyslu metody nejmenších čtverců pak bude znamenat ortogonální projekci funkčního vektoru f do lineárního obalu $\mathcal{L}(\varphi_1,\ldots,\varphi_k)$ aproximačních funkcí $\varphi_1,\ldots,\varphi_k$. Rozdíl $f-f^*$ je pak ortogonální komponentou, přičemž její velikost $||f-f^*||$ je jakožto velikost ortogonální komponenty minimální mezi všemi $||f-g||$ pro $g\in\mathcal{L}(\varphi_1,\ldots,\varphi_k)$. Zapíšeme-li tuto podmínku i s druhými mocninami (což je vlastně vynechání odmocnin při vyčíslování $||f-g||=(f-g)*(f-g)$, máme

$$||f-f^*||^2\leq||f-g||^2$$

pro všechny $g\in\mathcal{L}(\varphi_1,\ldots,\varphi_k)$, odtud název metody nejmenších čtverců.

V dalším kroku pak přejdeme k tzv. diskrétní metodě nejmenších čtverců. Vycházíme z údajů měření, které nám dává poznat hodnoty nějaké funkce jen ve vybraných hodnotách $t_1,\ldots,t_n$. Základní ideou je pak převod takovéhoto modelu na spojitou (to je v předchozím odstavci diskutovanou) MNČ s tím, že aproximační funkce $\varphi_1,\ldots,\varphi_k$ se nahrazují n-rozměrnými číselnými vektory $(\varphi_1(t_1),\varphi_1(t_n)),\ldots,$ $\varphi_k(t_1),\varphi_k(t_n))$ a funkce vystupující v roli f (nezáleží jaká mimo body $t_1,\ldots,t_n$) se nahrazuje vektorem $(f(t_1),\ldots,f(t_n))$. Integrální skalární součin se pak v této diskretizaci nahrazuje ,,běžným skalárním součinem".

Formulujme nyní problém MNČ pro spojitý a diskrétní případ a nakonec s využitím našeho geometrického přístupu jednotně pro oba případy.

a) Je dána spojitá funkce f na intervalu $\langle a,b\rangle$ a aproximační spojité funkce $\varphi_1,\ldots,\varphi_k$ definované na témže intervalu. Nalezněte lineární kombinaci f^* funkcí $\varphi_1,\ldots,\varphi_k$ splňující $||f-f^*||^2\leq||f-g||^2$ pro všechny lineární kombinace $g=\Sigma_{i=1}^k c_i\varphi_i$ aproximačních funkcí.

b) Jsou změřeny hodnoty neznámé funkce f v n vstupech (časových okamžicích) $x_1,\ldots,x_n$ s hodnotami $f_1,\ldots,f_k$. Dále jsou zadány aproximační funkce $\varphi_1,\ldots,\varphi_k$.

Nalezněte optimální lineární kombinaci f^* neznámé funkce f tak, aby mezi všemi funkcemi tvaru $g = \Sigma_{i=1}^k c_i\varphi_i$ minimalizovala $||f-g||^2 = \Sigma_{j=1}^n (f(x_j)-g(x_j))^2 = ||f - \Sigma_{i=1}^k c_i g_i||^2 = \Sigma_{j=1}^n (f - \Sigma_{i=1}^k c_i\varphi_i(x_j))^2$.

c) Je dán vektor f a vektory $\varphi_1, \ldots, \varphi_k$. Nalezněte optimální lineární kombinaci $f^* = \Sigma_{i=1}^k c_i^*\varphi_i$ vektorů $\varphi_1, \ldots, \varphi_k$ minimalizující $(f - \Sigma_{i=1}^k c_i\varphi_i) * (f - \Sigma_{i=1}^k c_i\varphi_i) = ||f - \Sigma_{i=1}^k c_i\varphi_i||^2$.

Návod na hledání optimálních koeficientů c_i^* dává následující věte, jejíž předpoklady lze vyjádřit pomocí bodu c). Věta navíc říká, že optimální koeficienty jsou určeny jednoznačně.

THEOREM 1. *Je dán vektor f a vektory $\varphi_1, \ldots, \varphi_k$. Pak existuje právě jedna sada koeficientů $c_1^*, \ldots, c_k^*$ minimalizující $(f - \Sigma_{i=1}^k c_i\varphi_i) * (f - \Sigma_{i=1}^k c_i\varphi_i) = ||f - \Sigma_{i=1}^k c_i\varphi_i||^2$, které jsou řešením následujícího systému lineárních rovnic*

$$\begin{pmatrix} \varphi_1 * \varphi_1 & \varphi_2 * \varphi_1 \ldots & \ldots & \varphi_k * \varphi_1 & |f * \varphi_1 \\ \varphi_1 * \varphi_2 & \varphi_2 * \varphi_2 \ldots & \ldots & \varphi_k * \varphi_2 & |f * \varphi_2 \\ \ldots & \ldots\ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots\ldots & \ldots & \ldots & \ldots \\ \varphi_1 * \varphi_k & \varphi_2 * \varphi_k \ldots & \ldots & \varphi_k * \varphi_k & |f * \varphi_k \end{pmatrix}$$

Příklad. Periodický signál $f(x) = 1$ pro $x \in \langle 0, \pi\rangle$ a $f(x) = 0$ pro $x \in (0, 2\pi)$ aproximujte spojitou metodou nejmenších čtverců pomocí aproximačních funkcí $\varphi_1 = 1$, $\varphi_2 = \cos x$ a $\varphi_3 = \sin x$.

Máme $\varphi_1 * \varphi_1 = \int_0^{2\pi} 1 \cdot 1 dx = 2\pi$, $\varphi_1 * \varphi_2 = 0 = \varphi_1 * \varphi_3 = \varphi_2 * \varphi_3$, $\varphi_2 * \varphi_2 = \int_0^{2\pi} \cos^2 x dx = \pi$ a $\varphi_3 * \varphi_3 = \int_0^{2\pi} \sin^2 x dx = \pi$. Dále $f * \varphi_1 = \int_0^{2\pi} f(x)\varphi_1(x) = \int_0^{\pi} dx = \pi$, $f * \varphi_2 = \int_0^{2\pi} f(x)\varphi_2(x) = \int_0^{\pi} \cos x dx = 0$ a $f * \varphi_3 = \int_0^{2\pi} f(x)\varphi_3(x) = \int_0^{\pi} \sin x dx = 2$. Optimální odhad ve smyslu MNČ bude tvaru $c_1 \cdot 1 + c_2 \cdot \cos x + c_3 \cdot \sin x$, kde c_1, c_2, c_3 jsou jediným řešením systému

$$\begin{pmatrix} 2\pi & 0 & 0 & |\pi \\ 0 & \pi & 0 & |0 \\ 0 & 0 & \pi & |2 \end{pmatrix}.$$

Tedy $c_1^* = \frac{1}{2}$, $c_2^* = 0$ a $c_3^* = \frac{2}{\pi}$ a $f^*(x) = \frac{1}{2} + \frac{2}{\pi}\sin x$.

Příklad. Měřením byla při zahřívání v čase $t[min]$ naměřena teplota $T[C]$

$$\begin{pmatrix} t[min] & 0 & 1 & 2 & 3 \\ T[C] & 10 & 25 & 42 & 58 \end{pmatrix}.$$

Metodou MNČ najděte co nejpřesnější lineární odhad závislosti teploty závisející na čase (tzv. regresní přímku teploty v závislosti na čase).

Za φ_1 vezmeme konstantní jednotkovou funkci a za φ_2 identickou funkci, tedy $\varphi 1(x) = 1$, $\varphi 2(x) = x$. Provedeme-li diskretizaci, máme $\varphi_1 \simeq (1,1,1,1)$ a $\varphi_2 \simeq (0,1,2,3)$. Dále $f \simeq (10, 25, 42, 58)$. Máme $\varphi_1 * \varphi_1 = 4$, $\varphi_1 * \varphi_2 = 6$, $\varphi_2 * \varphi_2 = 14$, $f * \varphi_1 = 135$ a $f * \varphi_2 = 283$. Výsledný odhad (regresní přímka má rovnici) $c_1^* + c_2^* x$, kde c_1, c_2 jsou jediným řešením systému lineárních rovnic

$$\begin{pmatrix} 4 & 6 & |135 \\ 6 & 14 & |283 \end{pmatrix}.$$

Výsledná funkce je tedy $T = 7,85 + 16,1 \cdot t$.

Poznámka. Aproximace polynomy 1. stupně v diskrétní metodě nejmenších čtverců odpovídá regresní přímce, zatímco aproximace konstantním polynomem (konstantou) odpovídá aritmetickému průměru ($\varphi_1 = 1 \simeq (1, \ldots, 1) \in \mathbb{R}^n$, tedy $\varphi_1 * \varphi_1 = n$. Dále $f * \varphi_1 \simeq f_1 \cdot 1 + \cdots + f_n \cdot 1$).

Příklad. Měřením bylo zjištěno, že molekulové hmotnosti oxidů dusíku jsou následující

$$\begin{pmatrix} N_2O & NO_2 & N_2O_3 & N_2O_5 \\ 44,013 & 46,006 & 76,012 & 108,010 \end{pmatrix}.$$

Určete atomové hmotnosti dusíku a kyslíku s co nejmenší chybou ve smyslu MNČ.

Označme atomovou hmotnost dusíku x a kyslíku y. Dostáváme následující systém lineárních rovnic v proměnných x, y.

$$\begin{pmatrix} 2 & 1 & |44,013 \\ 1 & 2 & |46,006 \\ 2 & 3 & |76,012 \\ 2 & 5 & |108,010 \end{pmatrix}.$$

Tento systém je tzv. přeurčený, což znamená, že je více rovnic než neznámých. Pokud jsme neměřili naprosto bezchybně (což je prakticky vyloučeno), systém nebude řešitelný. Správné hodnoty však lze odhadnout s co nejmenší chybou ve smyslu MNČ následovně. Přepišme systém jako

$$x \begin{pmatrix} 2 \\ 1 \\ 2 \\ 2 \end{pmatrix} + y \begin{pmatrix} 1 \\ 2 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 44,013 \\ 46,006 \\ 76,012 \\ 108,010 \end{pmatrix}.$$

První sloupcový vektor odpovídá φ_1, druhý φ_2 a třetí f. Máme $\varphi_1 * \varphi_1 = 13$, $\varphi_1 * \varphi_2 = 20$, $\varphi_2 * \varphi_2 = 39$, $f * \varphi_1 = 502,076$ a $f * \varphi_2 = 894,111$. Dostáváme systém lineárních rovnic, jehož řešením x, y jsou atomové hmotnosti dusíku a kyslíku

$$\begin{pmatrix} 13 & 20 & |502,076 \\ 20 & 39 & |894,111 \end{pmatrix}.$$

Poznamenejme, že funkce φ_1, φ_2 zde nejsou známy vůbec, známe je jen v hodnotách $1, 2, 3, 4$, přičemž tato čísla značí jen pořadí oxidu v tabulce. Jiným hodnotám "vstupů" asi těžko přiřadíme nějaký rozumný význam.