

Příklad (3.1)

Uvažujme neznámou funkci f , pro niž jsou nám známy její funkční hodnoty $f(x_i)$ v bodech $x_i, i \in [n]$. Tuto funkci se budeme snažit aproximovat polynomem p nejvýše k -tého stupně tak, aby

$$p = \arg \min_{\deg(q) \leq k} \sum_{i=1}^n (f(x_i) - q(x_i))^2.$$

Formulujte tuto rovnost jako problém nejmenších čtverců ($\arg \min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\|$) pro vhodnou matici A a vektory \mathbf{b} a \mathbf{x} .

┌

Řešení

Jelikož je q stupně nejvýše k , lze ho napsat jako: $q(y) = q_0 + q_1 y + q_2 y^2 + \dots + q_k y^k = \mathbf{q} \cdot \mathbf{y}$, kde $\mathbf{q} = (q_0, \dots, q_k)^T$ a $\mathbf{y} = (y^0, \dots, y^k)^T$. Rovnost výše pak můžeme přepsat na

$$\mathbf{p} = \arg \min_{\mathbf{q}} \sum_{i=1}^n (f_i - \mathbf{q} \cdot \mathbf{x}_i)^2.$$

Nyní se můžeme podívat z druhé strany: na vektorech uvažujeme l^2 normu, tedy $\arg \min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\|$ zapíšeme jako $\arg \min_{\mathbf{x}} \sqrt{\sum_{i=1}^n (\mathbf{a}_{i*} \cdot \mathbf{x} - b_i)^2}$. (Odmocnina ze součtu druhých mocnin jednotlivých složek.) Jelikož je odmocnina rostoucí funkce, bude minimální tam, kde bude minimální její argument, tedy je to totéž, co $\arg \min_{\mathbf{x}} \sum_{i=1}^n (\mathbf{a}_{i*} \cdot \mathbf{x} - b_i)^2$. (Jelikož výraz pod odmocninou je pro reálná čísla (na komplexních by úloha nefungovala) vždy kladný, nezvětšili jsme tím množinu vektorů, přes které hledáme minimum, tedy je to opravdu ekvivalentní.)

Když do této rovnice dosadíme $\mathbf{a}_{i*} = \mathbf{x}_i$, $\mathbf{x} = \mathbf{q}$ a $f_i = b_i$, tak dostaneme přesně pravou stranu rovnosti pro \mathbf{p} , tedy

$$\mathbf{p} = \arg \min_{\mathbf{x}} \sum_{i=1}^n (\mathbf{a}_{i*} \cdot \mathbf{x} - b_i)^2 = \arg \min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\|,$$

kde \mathbf{b} je vektor hodnot f v bodech x_i , A má za řádky řádkové vektory mocnin x_i a \mathbf{x} je vektor koeficientů hledaného polynomu.

└

Příklad (3.2)

Definujme $g(\mathbf{x}) = \frac{1}{2} \|A\mathbf{x} - \mathbf{b}\|^2$ a nahlížejme nyní na problém nejmenších čtverců jako na hledání extrému funkce více proměnných. Ukažte, že $\nabla_{\mathbf{x}} g(\mathbf{x}) = A^T(A\mathbf{x} - \mathbf{b})$. Diskutujte vztah hledání minima funkce g a řešení normálních rovnic.

┌ Důkaz (Vztahu pro ∇)

Přepíšeme g pomocí definice l^2 normy (výraz, který je v normě odmocňován je kladný, tedy $\sqrt{}$ a 2 se „požerou“ beze zbytku, n je dimenze prostoru, ve kterém problém řešíme, neboli třeba řád matice A):

$$g(\mathbf{x}) := \frac{1}{2} \|A\mathbf{x} - \mathbf{b}\|^2 = \frac{1}{2} \sum_{i=1}^n (\mathbf{a}_{i*} \cdot \mathbf{x} - b_i)^2 = \frac{1}{2} \sum_{i=1}^n (a_{i1}x_1 + \dots + a_{in}x_n - b_i)^2.$$

Parciální derivaci spočítáme tak, že pohlížíme na ostatní proměnné jako na konstanty, tedy z derivace složené funkce dostaneme:

$$\frac{\partial g}{\partial x_j}(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^n 2(a_{i1}x_1 + \dots + a_{in}x_n - b_i) \cdot (0 + \dots + a_{ij} + \dots 0) = \sum_{i=1}^n (\mathbf{a}_{i*}\mathbf{x} - b_i)a_{ij},$$

$$\frac{\partial g}{\partial x_j}(\mathbf{x}) = \mathbf{a}_{*j}(A\mathbf{x} - \mathbf{b}),$$

$$\nabla_{\mathbf{x}}g(\mathbf{x}) = (\mathbf{a}_{*1}(A\mathbf{x} - \mathbf{b}), \dots, \mathbf{a}_{*n}(A\mathbf{x} - \mathbf{b}))^T = A^T(A\mathbf{x} - \mathbf{b}).$$

□

┌ Řešení (Diskuze)

Z prvního příkladu víme, že $\arg \min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\| = \arg \min_{\mathbf{x}} \frac{1}{2} \|A\mathbf{x} - \mathbf{b}\|^2$ (funkce $\varphi(\mathbf{x}) = \frac{1}{2}\mathbf{x}$ je taktéž rostoucí), tedy hledání řešení problému nejmenších čtverců odpovídá hledání $\arg \min_{\mathbf{x}} g(\mathbf{x})$.

Z nutné podmínky existence lokálního extrému víme, že pokud je \mathbf{x} extrém g (a existují zde parciální derivace, což existují, když jsme je našli), tak

$$\nabla_{\mathbf{x}}g(\mathbf{x}) = A^T(A\mathbf{x} - \mathbf{b}) = \mathbf{0}.$$

Další extrémy mohou být už jen v ∞ , ale tam je $g(\infty) = +\infty$. Tedy všechna \mathbf{x} , pro která je $g(\mathbf{x})$ minimální, musí být řešením normální soustavy rovnic.

Zároveň $A^T(A\mathbf{x} - \mathbf{b}) = \mathbf{0}$ nám říká, že řádky A^T , tedy sloupce A jsou kolmé na $(A\mathbf{x} - \mathbf{b})$, tudíž z Pythagorovy věty (pro libovolné \mathbf{h})

$$\begin{aligned} g(\mathbf{x} + \mathbf{h}) &= \frac{1}{2} \|A(\mathbf{x} + \mathbf{h}) - \mathbf{b}\|^2 = \frac{1}{2} \|A\mathbf{h} + (A\mathbf{x} - \mathbf{b})\|^2 \\ &= \frac{1}{2} (\|A\mathbf{h}\|^2 + \|A\mathbf{x} - \mathbf{b}\|^2) \geq \frac{1}{2} \|A\mathbf{x} - \mathbf{b}\|^2 = g(\mathbf{x}). \end{aligned}$$

Tedy libovolné řešení normální soustavy rovnic je zároveň $\arg \min_{\mathbf{x}} g(\mathbf{x})$.

┌ Tj. řešení problému nejmenších čtverců jsou právě řešení normální soustavy rovnic.