Numerikus módszerek 2.

10. előadás: Az általánosított inverz és approximációs tulajdonsága

Krebsz Anna

ELTE IK

Tartalomjegyzék

Szemléltetés:

- Szimmetrikus A mátrix esetén $\exists$ U unitér és D diagonális mátrix, melyre $A = UDU^* \quad \Leftrightarrow \quad AU = UD$.
  Ez szemléletesen azt jelenti, hogy az A mátrix az $u_1, \ldots, u_n$ ONR-t a $\lambda_1 u_1, \ldots, \lambda_n u_n$ vektorokba viszi.
  A kép vektorok ortogonalitása továbbra is megmarad.

- Nem szimmetrikus esetben ez nem lehetséges. Azt szeretnénk elérni, hogy A az értelmezési tartomány ($\mathbb{R}^n$) $v_1, \ldots, v_r$ ortonormált vektorait (ahol $r = rang(A)$) a képhalmaz ($\mathbb{R}^m$) $\sigma_1 u_1, \ldots, \sigma_r u_r$ ortogonális vektoraiba vigye. Ekkor

$$A \cdot \begin{bmatrix} \vdots & \vdots & \vdots \\ v_1 & & v_r \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots \\ u_1 & & u_r \\ \vdots & \vdots & \vdots \end{bmatrix} \cdot \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix}.$$

Tétel: Szinguláris felbontás

Tetszőleges $A \in \mathbb{C}^{m\times n}$ esetén $\exists\ U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$ unitér mátrix és $D \in \mathbb{C}^{m\times n}$, particionálva $D = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}$, hogy

$$A = UDV^*,$$

ahol $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$, $r = \operatorname{rang}(A)$ és $\sigma_1 \geq \ldots \geq \sigma_r > 0$.

Definíció: Szinguláris értékek

A szinguláris felbontásbeli $\sigma_1, \ldots, \sigma_r > 0$ értékeket az A *szinguláris értékeinek* nevezzük.

$$\sigma_i = \sqrt{\lambda_i(A^*A)} > 0, \quad (i = 1, \ldots, r)$$

Megjegyzések:

- Az $A^*A \in \mathbb{C}^{n\times n}$ és $AA^* \in \mathbb{C}^{m\times m}$ pozitív sajátértékei azonosak.
- A^*A sajátvektorai a V oszlopai és
- AA^* sajátvektorai az U oszlopai.
- Nagyméretű, ritka A esetén a szinguláris értékeket nem A^*A-ból számítják, hanem az $\begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}$ particionált alak pozitív sajátértékeit határozzák meg, ezek a szinguláris értékek. Az A szinguláris vektorpárjait a sajátvektorok particionálásával kapják. (Hf.)

Speciális esetben: $m = n$ esetén

- $\sigma_1 = \max_{i=1}^n \sigma_i = \|A\|_2 = \|D\|_2$,
- $\|A\|_F^2 = \|D\|_F = \sum_{i=1}^n \sigma_i^2$.

Biz.:

- $A^*A \in \mathbb{C}^{n\times n}$ önadjungált (lásd előző félév) és pozitív szemidefinit, ugyanis
$$\langle A^*Ax, x\rangle = x^*(A^*Ax) = (Ax)^*Ax = \langle Ax, Ax\rangle \geq 0.$$
- A^*A pozitív sajátértékeit jelöljük $\sigma_1^2, \ldots, \sigma_r^2$-tel, továbbá a 0 $(n-r)$-szeres multiplicitású sajátérték.
$r = \text{rang}\,(A)$ és képezzük a $\Sigma = \text{diag}\,(\sigma_1, \ldots, \sigma_r) \in \mathbb{C}^{r\times r}$ mátrixot.
- A^*A önadjungált $\Rightarrow$ $\exists\ V \in \mathbb{C}^{n\times n}$ unitér: $A^*A = V\,D^2\,V^*$,
ahol $D^2 = \text{diag}\,(\lambda_i(A^*A)) = \begin{bmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{C}^{n\times n}$.
$A^*A \cdot V = V \cdot D^2$, vagyis V oszlopai az A^*A sajátvektorai.

Szinguláris felbontás bizonyítása

- Ha $n = m$ és $\operatorname{rang}(A) = n$, akkor A^*A pozitív definit, így

$$V^*A^*A\,V = D^2 \quad \Leftrightarrow \quad \underbrace{D^{-1}V^*A^*}_{=:U^*}\underbrace{A\,VD^{-1}}_{=:U} = I,$$

  ekkor ilyen egyszerű lenne a bizonyítás. Általában nem lehet D-t invertálni a 0 sajátértékek miatt, ezt kell kikerülnünk particionálással.

- Bontsuk V-t két részre: $V = [\underbrace{V_1}_{r}, \underbrace{V_2}_{n-r}]$, ahol V_1 tartalmazza a nem nulla sajátértékekhez tartozó sajátvektorokat és V_2 a 0 sajátértékekhez tartozó lin. fgt. sajátvektorokat.

- Külön vesszük a nem nulla sajátértékekhez tartozó részt:

$$V_1^* A^* A\, V_1 = \Sigma^2$$

$$\underbrace{\Sigma^{-1} V_1^* A^*}_{=:U_1^*} \underbrace{A\, V_1 \Sigma^{-1}}_{=:U_1} = I_{r\times r} \quad\Rightarrow\quad U_1 := A\, V_1 \Sigma^{-1}, \;\; U_1^* U_1 = I_{r\times r}$$

- U_1-et kiegészítjük $m - r$ db ortonormált $\mathbb{R}^m$-beli vektorral, ezek alkotják U_2-t. $U := [U_1, U_2]$, melyre $U^* U = I_{m\times m}$.

- Ellenőrizzük, hogy a kapott V és U unitér mátrixok jók:

$$U^* A\, V = \begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix} \cdot A \cdot [V_1, V_2] = \begin{bmatrix} U_1^* A V_1 & U_1^* A V_2 \\ U_2^* A V_1 & U_2^* A V_2 \end{bmatrix} = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}_{m\times n}.$$

  U_1 megadásából $AV_1 = U_1 \Sigma$ és U_1 és U_2 ortogonalitásából $U_2^*(AV_1) = U_2^*(U_1 \Sigma) = 0$, illetve $AV_2 = 0 \cdot V_2 = 0$ miatt a 2. oszlop elemei 0-k. $\square$

Alkalmazási területei:

1. Többváltozós statisztikában a szóródási mátrix vizsgálata, a variancia tömörítése.
2. Nagy dokumentumhalmazokban való keresés és tematizálás (Főkomponens analízis, adatbányászat).
3. Geofizikai mérések elemzésénél.

Tétel: Eckart–Young-tétel

Ha $\text{rang}\,(A) = r$, akkor a k rangú legjobb közelítése előállítható a szinguláris felbontásból

$$A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T,$$

ahol σ_i az i. szinguláris értéket, u_i és v_i a bal és jobb szinguláris vektort jelöli. Ez a közelítés a 2-es és a Frobenius normában is a legjobb, a hibák:

$$\|A - A_k\|_F = \sqrt{\sum_{i=k+1}^{r} \sigma_i^2}$$

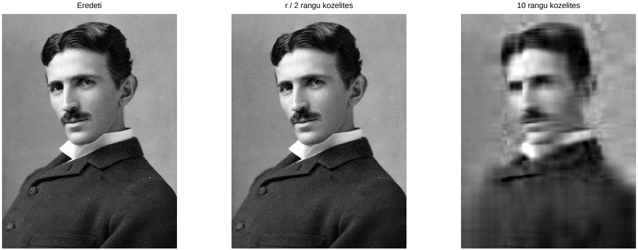
$$\|A - A_k\|_2 = \sigma_{k+1}.$$

Példa: Beadható Hf. villamosmérnököknek (BME):

Vegyünk egy pixelformátumú szürkeárnyalatos fényképet és készítsünk belőle egy mátrixot, melynek minden eleme a kép egy pontjának árnyalatát adja meg. Legyen a mátrix rangja r.

- Írjuk fel a mátrix szinguláris felbontását,
- majd abból készítsünk két olyan kép-mátrixot, hogy egyiknek $r/2$, másiknak 10 a rangja, és Frobenius-normában a legközelebb vannak az eredeti kép mátrixához.
- Határozzuk meg az eredetitől való eltérés Frobenius- és 2-normáját.

Alkalmazás jelfeldolgozásban



Definíció: Általánosított inverz

Az $A \in \mathbb{C}^{m \times n}$ mátrix *Moore-Penrose-féle általánosított (pszeudo) inverze* az $A^+ \in \mathbb{C}^{n \times m}$ mátrix, ha

1. AA^+ önadjungált,
2. A^+A önadjungált,
3. $AA^+A = A$,
4. $A^+AA^+ = A^+$.

Tétel:

A^+ egyértelmű.

Biz.: Lásd Csörgő István: Fejezetek a lineáris algebra köréből című könyvében.

Tétel: Diagonális mátrix általánosított inverze

A $D \in \mathbb{C}^{m\times n}$ diagonális mátrix általánosított inverze a $D^+ \in \mathbb{C}^{n\times m}$ diagonális mátrix, ahol $d_{ii}^+ = \frac{1}{d_{ii}}$, ha $d_{ii} \neq 0$, különben 0 az értéke.

Biz.: Az 1)-4) tulajdonságokat kell ellenőrizni.

Tétel: Az általánosított inverz előállítása

A szinguláris felbontás felhasználásával az $A \in \mathbb{C}^{m\times n}$ mátrix általánosított inverze a következő alakban állítható elő:

$$A^+ = VD^+U^*, \text{ ahol } D^+ \in \mathbb{C}^{n\times m} \text{ lásd az előző tételben.}$$

Biz.: Következik az előző tételből, az 1)-4) tulajdonságokat kell ellenőrizni.

Az általánosított inverz előállítása

Tétel: Túlhatározott teljes rangú eset

Legyen $A \in \mathbb{C}^{m \times n}$, $m > n$ és $r = \operatorname{rang}(A) = n$. Ekkor

$$A^+ = (A^*A)^{-1}A^*.$$

Biz.: Az 1)-4) tulajdonságokat kell ellenőrizni.

Tétel: Alulhatározott teljes rangú eset

Legyen $A \in \mathbb{C}^{m \times n}$, $m < n$ és $r = \operatorname{rang}(A) = m$. Ekkor

$$A^+ = A^*(AA^*)^{-1}.$$

Biz.: Az 1)-4) tulajdonságokat kell ellenőrizni.

1. Példa: általánosított inverzre

Készítsük el az alábbi túlhatározott teljes rangú mátrix általánosított inverzét.

$$A = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

Megoldás: Mivel a mátrix teljes rangú ($\operatorname{rang}(A) = 1$), túlhatározott esetben az általánosított inverz egyszerű képlettel számolható ($A^+ = (A^T A)^{-1} A^T$), nincs szükség a szinguláris felbontásra.

$$(A^T A) = \begin{bmatrix} 1 & 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} = 21, \quad (A^T A)^{-1} = \frac{1}{21}$$

$$A^+ = \frac{1}{21} \begin{bmatrix} 1 & 2 & 4 \end{bmatrix}$$

2. Példa: általánosított inverzre

Készítsük el a következő alulhatározott teljes rangú mátrix általánosított inverzét: $A = \begin{bmatrix} 1 & 2 & 4 \end{bmatrix}$.

Megoldás: Mivel a mátrix teljes rangú ($\operatorname{rang}(A) = 1$), alulhatározott esetben az általánosított inverz egyszerű képlettel számolható ($A^+ = A^T(AA^T)^{-1}$), nincs szükség a szinguláris felbontásra.

$$(AA^T) = \begin{bmatrix} 1 & 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} = 21, \quad (AA^T)^{-1} = \frac{1}{21}$$

$$A^+ = \frac{1}{21} \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

$A \in \mathbb{R}^{m\times n}$ esetben $A^T A$ általában nem invertálható, ilyenkor a túlhatározott esetre vonatkozó képlet nem alkalmazható, de kis módosítással az általánosított inverz közelíthető.

Tétel: Tyihonov-féle regularizáció

Legyen $A \in \mathbb{R}^{m\times n}$ és $m > n$, ekkor

$$\lim_{\varepsilon\to 0} (A^T A + \varepsilon^2 \cdot I_n)^{-1} A^T = A^+.$$

Biz.: Induljunk ki az $A = UDV^T$ szinguláris felbontásból:

$$\begin{aligned} A^T A + \varepsilon^2 \cdot I_n &= (UDV^T)^T UDV^T + \varepsilon^2 \cdot I_n \\ &= VD^T DV^T + \varepsilon^2 \cdot I_n = \\ &= V(D^T D + \varepsilon^2 \cdot I_n)V^T. \end{aligned}$$

A Tyihonov-féle regularizáció bizonyítása

Ennek inverze: $V(D^TD + \varepsilon^2 \cdot I_n)^{-1}V^T$.

Szorozzuk meg jobbról A^T-tal:

$$V(D^TD + \varepsilon^2 \cdot I_n)^{-1}V^T(UDV^T)^T = V(D^TD + \varepsilon^2 \cdot I_n)^{-1}D^TU^T$$

$(D^TD + \varepsilon^2 \cdot I_n)^{-1}D^T$ bal felső része:

$$(\Sigma^2 + \varepsilon^2 \cdot I_n)^{-1}\Sigma^T \to \Sigma^{-1} \ \ (\varepsilon \to 0).$$

Innen $(D^TD + \varepsilon^2 \cdot I_n)^{-1}D^T \to D^+ \ \ (\varepsilon \to 0)$.

A keresett határértékre

$$V\left[(D^TD + \varepsilon^2 \cdot I_n)^{-1}D^T\right]U^T \to VD^+U^T \ \ (\varepsilon \to 0).$$

□

Tartalomjegyzék

Az általánosított inverz approximációs tulajdonsága

Definíció: Általánosított megoldás

Legyen $A \in \mathbb{C}^{m\times n}$, $b \in \mathbb{C}^m$, ekkor az $x^+ := A^+ b \in \mathbb{C}^n$ vektort az $Ax = b$ *LER általánosított megoldásának* nevezzük.

Következmények:

- Ha A túlhatározott és teljes rangú, akkor

$$x^+ = A^+ b = (A^* A)^{-1} A^* b \iff A^* A\, x^+ = A^* b.$$

  Az $A^* A\, x = A^* b$ LER-t Gauss-féle normálegyenleteknek nevezik, x^+ a megoldása.

- Ha A alulhatározott és teljes rangú, akkor

$$x^+ = A^+ b = A^* \underbrace{(AA^*)^{-1} b}_{=:y} \iff AA^* y = b \text{ és } x^+ = A^* y.$$

- A fenti két esetben nem szükséges szinguláris felbontás az általánosított inverz előállításához.

Az általánosított inverz approximációs tulajdonsága

Tétel: Az általánosított inverz approximációs tulajdonsága

1. $\|Ax - b\|_2 \geq \|Ax^+ - b\|_2, \quad \forall\, x \in \mathbb{C}^n$.
2. $H := \{x \in \mathbb{C}^n |\ \|Ax - b\|_2 = \|Ax^+ - b\|_2\}$, akkor
$$\|x^+\|_2 < \|x\|_2 \ \forall\, x \in H, \ \ x \neq x^+.$$

Lemma:

1. Ha $x \perp y$ (azaz $\langle x, y \rangle = 0$), akkor $\|x + y\|_2^2 = \|x\|_2^2 + \|y\|_2^2$.
2. $A^*(AA^+ - I) = 0$
3. $(A^+)^*(I - A^+A) = 0$

Az általánosított inverz approximációs tulajdonsága

Biz.: Lemma:

1. Trivi, Pithagorasz tétele.
2. Az általánosított inverz 1) és 3) tulajdonságát felhasználva

$$\left(A^*(AA^+ - I)\right)^* = (AA^+ - I)^*A = ((AA^+)^* - I)A =_{1)}$$
$$= (AA^+ - I)A = AA^+A - A =_{3)} 0.$$

3. Az általánosított inverz 2) és 4) tulajdonságát felhasználva

$$\left((A^+)^*(I - A^+A)\right)^* = (I - A^+A)^*A^+ = (I - (A^+A)^*)A^+ =_{2)}$$
$$= (I - A^+A)A^+ = A^+ - A^+AA^+ =_{4)} 0.$$

Az általánosított inverz approximációs tulajdonsága

Tétel biz. 1.:

$$Ax - b = (Ax - Ax^+) + (Ax^+ - b) = A(x - x^+) + (Ax^+ - b)$$

Igazoljuk a 2. Lemmával, hogy a felbontás két komponense merőleges egymásra.

$$\langle A(x - x^+), Ax^+ - b \rangle = \langle x - x^+, A^*(Ax^+ - b) \rangle =$$
$$\langle x - x^+, A^*(AA^+ b - b) \rangle = \left\langle x - x^+, \underbrace{A^*(AA^+ - I)}_{=0}\, b \right\rangle = 0$$

Az 1. Lemmát felhasználva

$$\|Ax - b\|_2^2 = \|A(x - x^+) + (Ax^+ - b)\|_2^2 =$$
$$\underbrace{\|A(x - x^+)\|_2^2}_{\geq 0} + \|Ax^+ - b\|_2^2 \geq \|Ax^+ - b\|_2^2.$$

Az általánosított inverz approximációs tulajdonsága

Következmény: $x \in H$-ra $\|A(x - x^+)\|_2 = 0 \;\; \Rightarrow \;\; A(x - x^+) = 0.$

$$A^+ \cdot | \quad Ax = Ax^+$$

$$A^+Ax = A^+Ax^+ = \underbrace{A^+AA^+}_{=A^+} b = A^+b = x^+$$

Tétel biz. 2.:
Legyen $x \in H, \;\; x \neq x^+$ és $x = (x - x^+) + x^+$. Igazoljuk, hogy a felbontás két komponense merőleges.
A következményt és a 3. Lemmát felhasználva

$$x - x^+ = x - A^+Ax = (I - A^+A)x.$$

$$\langle x - x^+, x^+ \rangle = \langle (I - A^+A)x, A^+b \rangle = \left\langle \underbrace{(A^+)^*(I - A^+A)}_{=0} x, b \right\rangle = 0$$

Újra az 1. Lemmát felhasználva

$$\|x\|_2^2 = \underbrace{\|x - x^+\|_2^2}_{>0} + \|x^+\|_2^2 > \|x^+\|_2^2 \;\; \forall\, x \in H, \;\; x \neq x^+.$$

$\square$

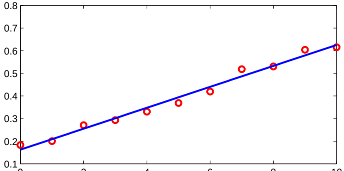
Tartalomjegyzék

Definíció: A legkisebb négyzetek módszerének alapfeladata

Adottak az $x_1, \ldots, x_N \in [a; b]$ különböző alappontok, $y_1, \ldots, y_N \in \mathbb{R}$ függvényértékek vagy mérési eredmények. Olyan $p_n \in P_n$ polinomot keresünk $(n + 1 \leq N$, általában $N \gg n)$, melyre

$$\sum_{i=1}^{N} (y_i - p_n(x_i))^2 \quad \text{minimális.}$$

A p_n polinomot *négyzetesen legjobban közelítő polinomnak* nevezzük.



Írjuk fel a $p_n(x_i) = y_i$, $(i = 1, \ldots, N)$ LER-t mátrix alakban, ahol

$$p_n(x) := a_n x^n + \ldots + a_1 x + a_0.$$

Vezessük be hozzá a következő jelöléseket:

$$A := \begin{bmatrix} 1 & x_1 & \ldots & x_1^n \\ 1 & x_2 & \ldots & x_2^n \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & x_N & \ldots & x_N^n \end{bmatrix}_{N \times (n+1)} \qquad a := \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}_{n+1 \times 1} \qquad y := \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_N \end{bmatrix}_{N \times 1} \qquad .$$

A kapott $A \cdot a = y$ LER klasszikus értelemben nem oldható meg, $N > n + 1$ esetén több egyenletünk van, mint ismeretlenünk. Különböző alappontok esetén $\operatorname{rang}(A) = n + 1$, vagyis teljes rangú LER-t kaptunk, amit általánosított értelemben meg tudunk oldani.

Túlhatározott teljes rangú esetben az általánosított megoldásra a Gauss-féle normálegyenleteket kapjuk:

$$a^+ = A^+ y = (A^T A)^{-1} \cdot A^T y \quad \Leftrightarrow \quad A^T A \cdot a^+ = A^T y.$$

A szimmetrikus LER alakja:

$$\begin{bmatrix} N & \sum x_i & \cdots & \sum x_i^n \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{n+1} \\ \vdots & & & \vdots \\ \sum x_i^n & \sum x_i^{n+1} & \cdots & \sum x_i^{2n} \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \vdots \\ \sum x_i^n y_i \end{bmatrix}.$$

Alkalmazzuk az approximációs tulajdonságot a LER-re:

$$\|A \cdot a^+ - y\|_2 \leq \|A \cdot a - y\|_2, \quad \forall\, a \in \mathbb{R}^{n+1}.$$

$$\|A \cdot a - y\|_2^2 = \sum_{i=1}^{N} (p_n(x_i) - y_i)^2 \;\; \rightarrow \;\; \text{minimalizálása},$$

így a^+ a négyzetesen legjobban közelítő polinom együtthatóit adja.

Megjegyzések:

- Ha A teljes rangú, akkor $A^T A$ mindig szimmetrikus és invertálható.
- A LER alakja $n = 1$ esetben:
$$\begin{bmatrix} N & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}.$$
- Ha $n = 1$ esetben $\sum x_i = 0$, akkor diagonális LER-t kapunk, melynek megoldása:
$$a_0 = \frac{1}{N} \sum y_i, \quad a_1 = \frac{\sum x_i y_i}{\sum x_i^2}.$$
A közgazdászok előszeretettel használják (lásd statisztika, regressziószámítás).
- A négyzetesen legjobban közelítő egyenes mindig átmegy az $\left(\frac{1}{N}\sum x_i, \frac{1}{N}\sum y_i\right)$ (átlagokból álló) ponton. ($n = 1$-re a Gauss-féle normálegyenletek első sora épp ezt jelenti.)

Legkisebb négyzetek módszere szélsőérték feladatként megoldva:
A négyzetesen legjobban közelítő polinomot $p_n(x) := a_n x^n + \ldots + a_1 x + a_0$ alakban keressük. Írjuk fel a minimalizálandó függvényt:

$$F(a_0, a_1, \ldots, a_n) := \sum_{i=1}^{N} (y_i - p_n(x_i))^2 = \sum_{i=1}^{N} \left(y_i - \sum_{j=0}^{n} a_j (x_i)^j \right)^2 .$$

A többváltozós valós függvény szélsőértékét keressük a derivált segítségével:

$$F'(a_0, a_1, \ldots, a_n) = 0$$

$$\frac{\partial F}{\partial a_k}(a_0, a_1, \ldots, a_n) = 0 \ \ (k = 0, \ldots, n).$$

$$k = 0, 1, \ldots, n : \qquad \frac{\partial F}{\partial a_k}(a_0, a_1, \ldots, a_n) = 0$$

$$\sum_{i=1}^{N} 2 \cdot (y_i - p_n(x_i)) \cdot \underbrace{\left(-\frac{\partial p_n}{\partial a_k}(x_i) \right)}_{-(x_i)^k} = 0$$

$$\sum_{i=1}^{N} 2 \cdot (y_i - p_n(x_i)) \cdot \left(-(x_i)^k \right) = 0 \quad | : 2$$

$$\sum_{i=1}^{N} p_n(x_i) \cdot (x_i)^k = \sum_{i=1}^{N} y_i (x_i)^k$$

$$\sum_{i=1}^{N} \sum_{j=0}^{n} a_j (x_i)^j \cdot (x_i)^k = \sum_{j=0}^{n} \sum_{i=1}^{N} a_j (x_i)^{j+k} = \sum_{i=1}^{N} y_i (x_i)^k$$

A belső szummából a_j-t kiemelve:

$$\sum_{j=0}^{n} a_j \sum_{i=1}^{N} (x_i)^{j+k} = \sum_{i=1}^{N} y_i (x_i)^k$$

A kapott LER a Gauss-féle normálegyenletek:

$$\begin{bmatrix} N & \sum x_i & \cdots & \sum x_i^n \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{n+1} \\ \vdots & & & \vdots \\ \sum x_i^n & \sum x_i^{n+1} & \cdots & \sum x_i^{2n} \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \vdots \\ \sum x_i^n y_i \end{bmatrix}.$$

Belátható, hogy a kapott $(a_0, a_1, \ldots, a_n)$ esetén $F''(a_0, a_1, \ldots, a_n)$ pozitív definit mátrix, így minimum helyet kaptunk.

Példa: legkisebb négyzetek módszere

Az alábbi $(x_i; y_i)$, $(i = 1, \ldots, 5)$ mérési eredményekhez határozzuk meg a négyzetesen legjobban közelítő egyenest.

x_i	-4	-1	0	2	3
y_i	-5	1	2	4	8

Megoldás: A négyzetesen legjobban közelítő elsőfokú polinom $(P_(x) = p_1 x + p_0)$ kiszámításához a következő lineáris egyenletrendszert kell megoldani:

$$\begin{bmatrix} N & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \cdot \begin{bmatrix} p_0 \\ p_1 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}.$$

Mivel a kitűzött feladatban

$$N = 5, \sum x_i = 0, \sum x_i^2 = 30, \sum y_i = 10, \sum x_i y_i = 51,$$

ezért a lineáris egyenletrendszer a következő.

$$\begin{bmatrix} 5 & 0 \\ 0 & 30 \end{bmatrix} \cdot \begin{bmatrix} p_0 \\ p_1 \end{bmatrix} = \begin{bmatrix} 10 \\ 51 \end{bmatrix}$$

Megoldva az egyenletrendszert

$$p_0 = \frac{10}{5} = 2$$
$$p_1 = \frac{51}{30} = \frac{17}{10}$$

A keresett egyenes egyenlete $P_1(x) = 1,7\,x + 2$.

Köszönöm a figyelmet!