Ortogonalprojektion, Mindstekvadratersproblem, Afsnit 6.3 og 6.5

19. april 2021

Sandsynlighedsteori og lineær algebra (SLIAL)

Forår 2021

AALBORG UNIVERSITY
DENMARK

Part I

Repetition

Underrum, basis

$u, v \in W \Rightarrow u+v \in W$

$u \in W, \alpha \in \mathbb{R} \Rightarrow \alpha u \in W$

$W \subset \mathbb{R}^n$: $0 \in W$, lukket under addition og multiplikation med skalarer

$W = \text{span}(v_1, \ldots, v_k)$, hvor $v_1, \ldots, v_k$ er lineært uafhengige

Indre produkt, ortogonalitet

$u, v \in \mathbb{R}^n$

Indre produkt:

$$u \cdot v = u_1 v_1 + \ldots u_n v_n$$

Norm:

$$||u|| = \sqrt{u \cdot u}$$

Ortogonalitet:

$$u \cdot v = 0 \quad \Longleftrightarrow \quad u \perp v$$

Part II

Ortogonal projektion

Ortogonalprojektion

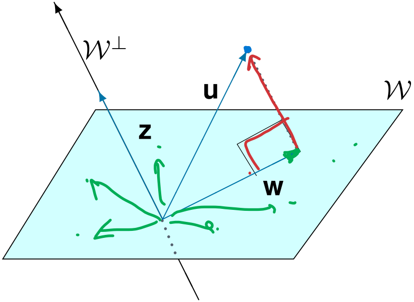
Givet: $\mathbf{u} \in \mathbb{R}^n$, $\mathcal{W} \subset \mathbb{R}^n$-underrum.

Vi ønsker at projektere en vektor $\mathbf{u}$ ned på underrummet W. Dvs. vi ønsker at bestemme den $\mathbf{w} \in \mathcal{W}$ som er tættest på $\mathbf{u}$:

$$\underline{\min_{\mathbf{w} \in \mathcal{W}} \|\mathbf{u} - \mathbf{w}\|^2.}$$

Notation: $\boxed{\mathbf{w} = \mathrm{proj}_{\mathcal{W}}(\mathbf{u})}$

Geometri af ortogonalprojektion



$$\mathbf{z} = \mathbf{u} - \mathbf{w} \in \mathcal{W}^{\perp}$$

hvor $\mathcal{W}^{\perp} = \{\, \mathbf{v} \in \mathbb{R}^n : \mathbf{v} \cdot \mathbf{x} = 0, \forall \mathbf{x} \in \cancel{\mathbb{R}^n} \, \mathcal{W} \,\}$ kaldes for et *ortogonal komplement* af $\mathcal{W}$.

Hvorfor er det sådan?

Antag at $\mathbf{w}$ er en projektion af $\mathbf{u}$ på $\mathcal{W}$, $\mathbf{z} = \mathbf{u} - \mathbf{w}$, $\mathbf{v} \in \mathcal{W}$, og $\alpha \in \mathbb{R}$.
Da $\mathbf{w} + \alpha\mathbf{v} \in \mathcal{W}$, og

$$\|\mathbf{u} - (\mathbf{w} + \alpha\mathbf{v})\|^2 = (\mathbf{z} - \alpha\mathbf{v}) \cdot (\mathbf{z} - \alpha\mathbf{v}) = \|\mathbf{z}\|^2 - 2\alpha\mathbf{z} \cdot \mathbf{v} + \alpha^2\|\mathbf{v}\|^2$$
$$\geq \underbrace{\|\mathbf{z}\|^2 = \|\mathbf{u} - \mathbf{w}\|^2}_{\text{mindst mulig!}}$$

$$\alpha(\alpha\|v\|^2 - 2z \cdot v) \geq 0$$

Dvs:

$$\boxed{\alpha^2\|\mathbf{v}\|^2 - 2\alpha\mathbf{z} \cdot \mathbf{v} \geq 0,} \quad \forall\alpha \in \mathbb{R} \text{ og } \mathbf{v} \in \mathcal{W}$$

Er kun mulig hvis $\mathbf{z} \in \mathcal{W}^\perp$

Hvornår er $\mathbf{z} \in \mathcal{W}^\perp$?

Antag at $\mathcal{W} = \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$, og $\mathbf{z} \in \mathcal{W}^\perp$.

Da $\mathbf{z} \cdot \mathbf{v}_i = 0$, $i = 1, \dots, k$.

Bevis:

$$v_i = 0 \cdot v_1 + 0 v_2 + \dots + 1 \cdot v_i + \dots + 0 \cdot v_k$$

$$\in \text{span}(v_1 \dots v_k) = W$$

$$z \in W^\perp \Rightarrow z \cdot v_i = 0$$

Hvornår er $\mathbf{z} \in \mathcal{W}^{\perp}$?

Antag at $\mathcal{W} = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, og $\mathbf{z} \cdot \mathbf{v}_i = 0$, $i = 1, \ldots, k$.

Da $\mathbf{z} \in \mathcal{W}^{\perp}$

Bevis: $v \in W = \text{span}(v_1 \ldots v_k)$

$$v = c_1 v_1 + c_2 v_2 + \ldots + c_k v_k$$

$$v \cdot z = c_1 \underbrace{v_1 \cdot z}_{=0} + c_2 \underbrace{v_2 \cdot z}_{=0} + \ldots + c_k \underbrace{v_k \cdot z}_{=0}$$

$$= 0 \quad \Rightarrow \quad z \in W^{\perp}$$

Beregning af projektion

Antag at $\mathcal{W} = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, $\mathbf{v}_i$-ON basis for $\mathcal{W}$.

$$\mathbf{z} = \mathbf{u} - \underbrace{\sum_{i=1}^{k} \alpha_i \mathbf{v}_i}$$

Vil bestemme $\alpha_i \in \mathbb{R}$: $\mathbf{z} \in \mathcal{W}^{\perp}$.

$$0 = z \cdot v_j = u \cdot v_j - \sum_{i=1}^{k} \alpha_i \underbrace{v_i \cdot v_j}_{\substack{=0,\ i \neq j \\ =1,\ i=j}}$$

$$= u \cdot v_j - \alpha_j = 0$$

$$\alpha_j = u \cdot v_j$$

Beregning af projektion

α_i

Konklusion:

$$\boxed{\mathbf{w} = \mathrm{proj}_{\mathcal{W}}(\mathbf{u}) = \sum_{i=1}^{k} (\mathbf{u} \cdot \mathbf{v}_i)\mathbf{v}_i,}$$

hvor $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ er ON-basis for $\mathcal{W}$.

Eksempel

$\in \mathbb{R}^4$

Lad $\mathbf{u} = [0.5, 0.5, 0.5, 0.5]$, og $\mathcal{W} = \text{span}([1, 2, 1, 2])$. Bestem $\text{proj}_{\mathcal{W}}(\mathbf{u})$.

$$v_1 = \frac{1}{\|[1,2,1,2]\|} \cdot \begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \end{bmatrix} = \frac{1}{\sqrt{1+4+1+4}} \cdot \begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \end{bmatrix}$$

$$w = (u \cdot v_1) v_1 = \left(\frac{1}{2} \cdot \frac{1}{\sqrt{10}} + \frac{1}{2} \cdot \frac{2}{\sqrt{10}} + \frac{1}{2} \cdot \frac{1}{\sqrt{10}} + \frac{1}{2} \cdot \frac{2}{\sqrt{10}} \right) \cdot v_1$$

$$= \left(\frac{6}{2\sqrt{10}} \right) \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix} = \frac{6}{20} \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix} = \frac{3}{10} \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}$$

$\text{proj}_W(u)$

Part III

Mindstekvadratersproblemer

Mindstekvadratersproblemer

Givet:

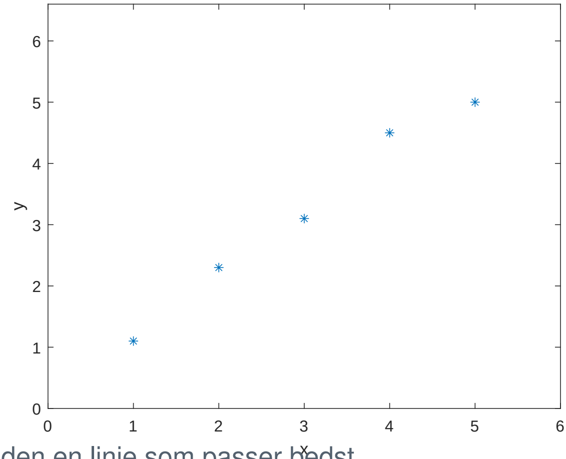
- $m \times n$ matrisa A
- $\mathbf{b} \in \mathbb{R}^m$

bestem $\hat{\mathbf{x}} \in \mathbb{R}^n$ således at

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| = \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{b} - A\mathbf{x}\|$$

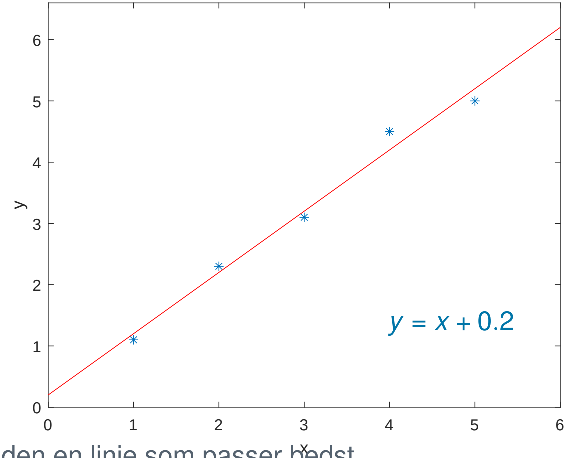
$\|Ax - b\| \to \min!$

Eksempel: Regression



Bestem den en linje som passer bedst.

Eksempel: Regression



Bestem den en linje som passer bedst.

Eksempel: Regression

Vi ønsker at bestemme det $n-1$'te grads polynomium som minimerer afstanden

$$\sum_{i=1}^{m}[p(t_i)-y_i]^2,$$

hvor t_i og y_i er givet, og

$$p(t) = a_0 + a_1 t + \cdots + a_{n-1} t^{n-1}.$$

Dvs, givet $\mathbf{b} = [y_1, \ldots, y_m] \in \mathbb{R}^m$, vi vil bestemme $\hat{\mathbf{x}} = [a_0, \ldots, a_{n-1}] \in \mathbb{R}^n$ således at polynomet passet til punkterne.

Eksempel: Regression

Lad

$$A = \begin{bmatrix} 1 & t_1 & t_1^2 & \dots & t_1^{n-1} \\ 1 & t_2 & t_2^2 & \dots & t_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & t_m & t_m^2 & \dots & t_m^{n-1} \end{bmatrix}$$

Da

$$A\mathbf{x} = A \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} p(t_1) \\ p(t_2) \\ \vdots \\ p(t_m) \end{bmatrix}$$

og

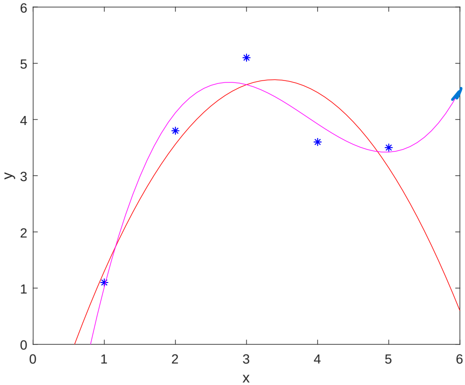
$$\|A\mathbf{x} - b\|^2 = \sum_{i=1}^{m} (p(t_i) - y_i)^2$$

Eksempel: Regression

De bedste polynome af grad 2 og 3 til punkter

$t_1 \quad y_1 \quad t_2 \quad y_2 \quad \cdots$

$$\{(1, 1.1), (2, 3.8), (3, 5.1), (4, 3.6), (5, 3.5)\}$$



Geometri av mindstekvadraters

Lad $\mathcal{W} = \text{span}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$, hvor $\mathbf{a}_i$ er søjler af A.

$w \in W$: $\|w - b\| \to \min$

$x \in \mathbb{R}^n$: $w = Ax$, $\|w - b\| \to \min$

$x \in \mathbb{R}^n$ $\|Ax - b\| \to \min!$

Vi vil da finne $\hat{\mathbf{b}} = \text{proj}_{\mathcal{W}}(\mathbf{b})$.

Fordi $\hat{\mathbf{b}} \in \mathcal{W} = \text{span}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$, finnes det $\hat{\mathbf{x}} \in \mathbb{R}^n$, således at $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$.

Denne $\hat{\mathbf{x}}$ er en løsning til mindstekvadratersproblemet

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|A\mathbf{x} - \mathbf{b}\|$$

Hvordan bestemmer vi $\hat{b} = A\hat{\mathbf{x}}$?

Vi ved at

$$\mathbf{z} = \mathbf{b} - \text{proj}_{\mathcal{W}}(\mathbf{b}) \in \mathcal{W}^{\perp}$$

$W = \text{Span}(a_1, \dots, a_n)$

.

Sker kun og hvis kun

$$\underline{\mathbf{z} \cdot \mathbf{a}_i = 0,} \qquad i = 1, \dots, n \qquad \Longleftrightarrow \qquad A^T\mathbf{z} = \mathbf{0}$$

Normalligningerne

$$A^T\mathbf{z} = A^T\underbrace{[\mathbf{b} - A\hat{\mathbf{x}}]}_{z} = \mathbf{0}.$$

Konklusion

Løsning til mindstekvadratersproblemet finnes som en løsning til ligningsystemet

$$\boxed{\underbrace{A^T A}_{n \times n}\hat{\mathbf{x}} = A^T\mathbf{b}} \iff z \in W^\perp$$

Eksempel: regression

Bestem den bedste parabel $y = a_0 + a_1 t + a_2 t^2$ til

$$\begin{array}{cc} t_i & y_i \\ \end{array}$$

$$\{(1, 1.1), (2, 3.8), (3, 5.1), (4, 3.6), (5, 3.5)\}$$

$$A = \begin{array}{c} \begin{array}{ccc} 1 & t_i & t_i^2 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{bmatrix} \end{array}, \quad \mathbf{b} = \begin{array}{c} y_i \\ \begin{bmatrix} 1.1 \\ 3.8 \\ 5.1 \\ 3.6 \\ 3.5 \end{bmatrix} \end{array}$$

og

$$A^\top A = \begin{bmatrix} 5 & 15 & 55 \\ 15 & 55 & 225 \\ 55 & 225 & 979 \end{bmatrix}, \quad A^\top \cancel{\mathbf{y}}\, b = \begin{bmatrix} 17.1 \\ 55.9 \\ 207.3 \end{bmatrix}$$

Eksempel: regression

Normalligningerne

$$\begin{bmatrix} 5 & 15 & 55 \\ 15 & 55 & 225 \\ 55 & 225 & 979 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 17.1 \\ 55.9 \\ 207.3 \end{bmatrix}$$

Løsningen er

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -2.16 \\ 4.06 \\ -0.6 \end{bmatrix}$$

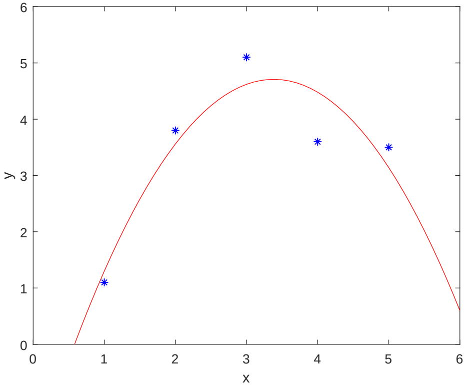
så

$$\boxed{y = -0.6t^2 + 4.06t - 2.16}$$

er det bedste fit.

Eksempel: regression

$$y = -0.6t^2 + 4.06t - 2.16$$



QR-faktorisering

Vi hadde sett at ortogonal projektion $\mathrm{proj}_{\mathcal{W}}$ er nemt å bestemme hvis vi kender ON basis til $\mathcal{W}$.

I tilfelle av mindstekvadraters er $\mathcal{W} = \mathrm{span}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ hvor $\mathbf{a}_i$ er søjler af A.

Lad matrisa Q indenholde en ON basis til $\mathcal{W}$ som søjler; dvs Q er *ortogonal*, $Q^T Q = I$.

Da

$$\hat{\mathbf{b}} = \mathrm{proj}_{\mathcal{W}}(\mathbf{b}) = \sum_{i=1}^{n} (\mathbf{q}_i \cdot \mathbf{b})\mathbf{q}_i = Q(Q^T \mathbf{b})$$

Skal "kun" finne $\hat{\mathbf{x}}$ således at $A\mathbf{x} = \hat{\mathbf{b}}$

QR-faktorisering

Antag nu at A kan skrives som QR, hvor

- Q er ortogonal, indenholder ON basis til søjlerum af A
- R er øvre triangulær, dvs, $R_{ij} = 0$, $i > j$.

Næste gang skal vi se på algoritmer, som beregner sådan en faktorisering.

Q^T $\quad$ Q^T $\quad$ $Q^TQ = I$

Da:

$$A\hat{\mathbf{x}} = QR\hat{\mathbf{x}} = \hat{\mathbf{b}} = Q(Q^T\mathbf{b})$$

$$R\hat{\mathbf{x}} = Q^T\mathbf{b}$$

QR-faktorisering

Den sidste ligning er nemt å løse (allerede i række-echelon formen)

$$R\hat{\mathbf{x}} = Q^T\mathbf{b}$$

Mindste kvadraters metode

Bestem mindste kvadraters løsning til ligningssysstemet

$$\underbrace{\begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 1 & 1 \end{bmatrix}}_{A} \mathbf{x} = \underbrace{\begin{bmatrix} 7 \\ 3 \\ 1 \end{bmatrix}}_{b}$$

Mindste kvadraters metode

Bestem mindste kvadraters løsning til ligningssysstemet

$$\begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 1 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 7 \\ 3 \\ 1 \end{bmatrix}$$

Ved at lave en QR-faktorisering af A fås at

$$A = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 1 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 3 & 5 \\ 0 & 1 \end{bmatrix}}_{R}$$

Mindste kvadraters metode

Vi løser derfor ligningssystemet

$$\overbrace{\begin{bmatrix} 3 & 5 \\ 0 & 1 \end{bmatrix}}^{R} \mathbf{x} = \overbrace{\begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{bmatrix}}^{Q^T} \begin{bmatrix} 7 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \end{bmatrix}$$

som har løsningen $x_2 = -1$ og

$$3x_1 + 5x_2 = 7 \Rightarrow x_1 = 4$$

Eksistens og entydighet af løsninger

Løsningen til mindstekvadratersproblemet eksisterer ($\mathbf{b}$ giver entydig $\hat{\mathbf{b}}$ i søjlerum af A, som betyder at mindst en $\hat{\mathbf{x}} \in \mathbb{R}^n$ opfylder $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$).

Løsningen er entydig, for alle $b \in \mathbb{R}^m \iff$ søjlene af A er lineært uafhengige $\iff A^T A$ er invertibel (se Theorem 14 i afnist 6.5).

$a_1 \ldots a_m$ - lin uafheng.

$\hat{x}^1, \hat{x}^2$ - løsn.

$A\hat{x}^1 = \hat{b}$

$A\hat{x}^2 = \hat{b}$ $\Rightarrow$ $A(\hat{x}^1 - \hat{x}^2) = 0$

$\underbrace{(\hat{x}^1_1 - \hat{x}^2_1)}_{=0} a_1 + \ldots + \underbrace{(\hat{x}^1_n - \hat{x}^2_n)}_{=0} a_n = 0$

$\Rightarrow \hat{x}^1 = \hat{x}^2$