Rezolvări Seminar

Cristian Rusu

Octombrie 2025

Problema 1

Potriviți dreptele $y_i = ax_i$ și $y_i = ax_i + b$ în sensul celor mai mici pătrate pentru perechile de puncte din setul de date $\{(x_i, y_i)\}$:

$$\{(0,1),\ (1,2),\ (2,2)\}. \tag{1}$$

Determinați a și (a, b) și suma pătratelor reziduurilor pentru:

$$\underset{a}{\text{minimizează}} \sum_{i=1}^{3} (y_i - ax_i)^2. \tag{2}$$

$$\underset{a,b}{\text{minimizează}} \sum_{i=1}^{3} (y_i - (ax_i + b))^2. \tag{3}$$

Soluție.

Vom nota vectorii:

$$\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}. \tag{4}$$

Problema noastră devine:

$$\underset{a}{\text{minimizează}}\ \|\mathbf{y} - \mathbf{ax}\|_2^2. \tag{5}$$

Soluția folosind ecuațiile normale este:

$$\begin{aligned} a^\star &= (\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\mathbf{y} = \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{x}^T\mathbf{x}} = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|_2^2} \\ &= \frac{1 \times 0 + 2 \times 1 + 2 \times 2}{0^2 + 1^1 + 2^2} = \frac{6}{5}; \end{aligned} \tag{6}$$

Erorile noastre sunt:

$$\mathbf{r} = \mathbf{y} - a^\star \mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{6}{5} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{4}{5} \\ -\frac{2}{5} \end{bmatrix} \tag{7}$$

Eroarea totală este:

$$\|\mathbf{r}\|_2^2 = 1^2 + \left(\frac{4}{5}\right)^2 + \left(-\frac{2}{5}\right)^2 = 1 + \frac{20}{25} = 1 + \frac{4}{5}. \tag{8}$$

Observați că $\mathbf{r}^T\mathbf{x} = 0$.

Vom nota vectorul și matricea:

$$\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}. \tag{9}$$

$$\underset{a,b}{\text{minimizează}} \ \|\mathbf{y} - \mathbf{X} \begin{bmatrix} a \\ b \end{bmatrix} \|_2^2. \tag{10}$$

Soluția folosind ecuațiile normale este:

$$\begin{aligned} \begin{bmatrix} a^\star \\ b^\star \end{bmatrix} &= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \\ &= \begin{bmatrix} 5 & 3 \\ 3 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 5 \end{bmatrix} \\ &= \frac{1}{6} \begin{bmatrix} 3 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} 6 \\ 5 \end{bmatrix} \\ &= \frac{1}{6} \begin{bmatrix} 3 \\ 7 \end{bmatrix} \\ &= \begin{bmatrix} 1/2 \\ 7/6 \end{bmatrix} \end{aligned} \tag{11}$$

Erorile noastre sunt:

$$\begin{aligned} \mathbf{r} = \mathbf{y} - \mathbf{X} \begin{bmatrix} a^\star \\ b^\star \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{1}{6} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 7 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{1}{6} \begin{bmatrix} 7 \\ 10 \\ 13 \end{bmatrix} \\ &= \begin{bmatrix} -1/6 \\ 1/3 \\ -1/6 \end{bmatrix} \end{aligned} \tag{12}$$

Eroarea totală este:

$$\|\mathbf{r}\|_2^2 = \left(-\frac{1}{6}\right)^2 + \left(\frac{2}{6}\right)^2 + \left(-\frac{1}{6}\right)^2 = \frac{1+4+1}{36} = \frac{6}{36} = \frac{1}{6}. \tag{13}$$

Observați că $\mathbf{r}^T\mathbf{X} = \mathbf{0}_{2\times 1}$, din nou. Observați și faptul că reziduul a scăzut semnificativ.

Problema 2

Se dau matricea $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ și vectorul $\mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ (de discutat putin de ce se schimbă notația).

1. Aplicați Gram-Schmidt pentru a scrie $\mathbf{A} = \mathbf{QR}$, cu $\mathbf{Q} \in \mathbb{R}^{3\times 2}$ având coloane ortonormate. Folosiți $\mathbf{QR}$ pentru a obține soluția de cele mai mici pătrate $\mathbf{x}^\star$.

2. Calculați factorizarea QR cu reflector Householder $\mathbf{A} = \mathbf{QR}$, cu $\mathbf{Q} \in \mathbb{R}^{3\times 2}$ având coloane ortonormate. Folosiți $\mathbf{QR}$ pentru a obține soluția de cele mai mici pătrate $\mathbf{x}^\star$.

Soluție. Aplicați Gram–Schmidt pe coloanele $\mathbf{a}_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix}1\\0\\1\end{bmatrix}$:

- $\mathbf{q}_1 = \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\end{bmatrix}$.
- Proiecția lui $\mathbf{a}_2$ pe $\mathbf{q}_1$:

$$(\mathbf{q}_1^T\mathbf{a}_2)\mathbf{q}_1 = \tfrac{1}{\sqrt{2}} \cdot \tfrac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\end{bmatrix} = \begin{bmatrix}1/2\\1/2\\0\end{bmatrix}. \tag{14}$$

$$\mathbf{u}_2 = \mathbf{a}_2 - (\mathbf{q}_1^T\mathbf{a}_2)\mathbf{q}_1 = \begin{bmatrix}1/2\\-1/2\\1\end{bmatrix}, \|\mathbf{u}_2\| = \frac{\sqrt{6}}{2}, \tag{15}$$

$$\mathbf{q}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = \frac{1}{\sqrt{6}}\begin{bmatrix}1\\-1\\2\end{bmatrix}. \tag{16}$$

Prin urmare

$$\mathbf{Q} = \begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}}\\ 0 & \frac{2}{\sqrt{6}}\end{bmatrix}, \mathbf{R} = \mathbf{Q}^T\mathbf{A} = \begin{bmatrix}\sqrt{2} & \frac{1}{\sqrt{2}}\\ 0 & \frac{\sqrt{6}}{2}\end{bmatrix}. \tag{17}$$

Soluția celor mai mici pătrate se obține din $\mathbf{R}\mathbf{x} = \mathbf{Q}^T\mathbf{b}$. Calculăm

$$\mathbf{Q}^T\mathbf{b} = \begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{6}}\end{bmatrix}. \tag{18}$$

Substituție inversă:

$$\frac{\sqrt{6}}{2}x_2 = \frac{1}{\sqrt{6}} \Rightarrow x_2 = \frac{1}{3}, \qquad \sqrt{2}\,x_1 + \frac{1}{\sqrt{2}}x_2 = \frac{1}{\sqrt{2}} \Rightarrow 2x_1 + \frac{1}{3} = 1 \Rightarrow x_1 = \frac{1}{3}. \tag{19}$$

$$\mathbf{x}^\star = \begin{bmatrix}1/3\\1/3\end{bmatrix}. \tag{20}$$

Realizăm factorizarea $\mathbf{A} = \mathbf{Q}\mathbf{R}$ cu reflectori Householder și apoi calculăm soluția celor mai mici pătrate din $\mathbf{R}\mathbf{x} = \mathbf{Q}^T\mathbf{b}$.

Pasul 1. Eliminăm toate elementele sub-diagonale de pe prima coloană. Pentru $\mathbf{a}_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}$, $\|\mathbf{a}_1\| = \sqrt{2}$. Alegem

$$\mathbf{v}_1 = \mathbf{a}_1 + \|\mathbf{a}_1\|\,\mathbf{e}_1 = \begin{bmatrix}1+\sqrt{2}\\1\\0\end{bmatrix}, \quad \mathbf{H}_1 = \mathbf{I} - 2\frac{\mathbf{v}_1\mathbf{v}_1^T}{\mathbf{v}_1^T\mathbf{v}_1}. \tag{21}$$

Atunci $\mathbf{H}_1\mathbf{a}_1 = -\|\mathbf{a}_1\|\mathbf{e}_1 = \begin{bmatrix}-\sqrt{2}\\0\\0\end{bmatrix}$, iar

$$\mathbf{H}_1\mathbf{A} = \begin{bmatrix}-\sqrt{2} & -\frac{\sqrt{2}}{2}\\ 0 & -\frac{\sqrt{2}}{2}\\ 0 & 1\end{bmatrix} = \mathbf{A}^{(1)}. \tag{22}$$

Pasul 2. Eliminăm toate elementele sub-diagonale de pe a doua coloană. Considerăm subvectorul din a doua coloană $\mathbf{z} = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ 1 \end{bmatrix} \in \mathbb{R}^2$, cu $\|\mathbf{z}\| = \sqrt{\frac{3}{2}} = \frac{\sqrt{6}}{2}$.

Luăm $\mathbf{v}_2 = \mathbf{z} - \alpha \mathbf{e}_1$, unde $\alpha = -\operatorname{sign}(z_1)\|\mathbf{z}\| = \frac{\sqrt{6}}{2}$, și reflectorul

$$\widehat{\mathbf{H}}_2 = \mathbf{I}_2 - 2\frac{\mathbf{v}_2\mathbf{v}_2^T}{\mathbf{v}_2^T\mathbf{v}_2} = \begin{bmatrix} -\frac{1}{\sqrt{3}} & \sqrt{\frac{2}{3}} \\ \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}. \tag{23}$$

Îl înglobăm în $\mathbb{R}^{3\times 3}$ și obținem:

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{\sqrt{3}} & \sqrt{\frac{2}{3}} \\ 0 & \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}. \tag{24}$$

Atunci $\mathbf{R} = \mathbf{H}_2\mathbf{A}^{(1)}$ este superior triunghiular (pe primii doi rânduri/coloane):

$$\mathbf{R} = \begin{bmatrix} -\sqrt{2} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{\sqrt{6}}{2} \\ 0 & 0 \end{bmatrix}. \tag{25}$$

Pentru că reflectorii sunt simetrici și ortogonali, avem $\mathbf{Q} = \mathbf{H}_1\mathbf{H}_2$, iar $\mathbf{Q}^T = \mathbf{H}_2\mathbf{H}_1$.

Pentru soluția problemei celor mai mici pătrate vom calcula.

$$\mathbf{y} = \mathbf{Q}^T\mathbf{b} = \mathbf{H}_2\mathbf{H}_1\mathbf{b} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} \\ ? \end{bmatrix}, \tag{26}$$

iar sistemul ce trebuie rezolvat este:

$$\begin{bmatrix} -\sqrt{2} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{\sqrt{6}}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}. \tag{27}$$

Din a doua ecuație: $\frac{\sqrt{6}}{2}x_2 = \frac{1}{\sqrt{6}} \Rightarrow x_2 = \frac{2}{6} = \frac{1}{3}$. Din prima: $-\sqrt{2}\,x_1 - \frac{1}{\sqrt{2}}x_2 = -\frac{1}{\sqrt{2}} \Rightarrow 2x_1 + x_2 = 1 \Rightarrow x_1 = \frac{1}{3}$.

$$\mathbf{x}^\star = \begin{bmatrix} 1/3 \\ 1/3 \end{bmatrix}. \tag{28}$$

Observație: semnele diagonalei lui $\mathbf{R}$ pot fi făcute pozitive printr-o alegere echivalentă a reflectorilor; soluția LS rămâne aceeași.

Problema 3

Pentru un $\mathbf{A} \in \mathbb{R}^{m\times n}$ general:

(a) Arătați că reziduul de celor mai mici pătrate $\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}^\star$ este ortogonal pe toate coloanele matricei $\mathbf{A}$, adică $\mathbf{A}^T\mathbf{r} = \mathbf{0}_{n\times 1}$.

(b) Calculați proiectorul ortogonal $\mathbf{P}$ pe coloanele matricei $\mathbf{A}$ și folosiți-l pentru a obține vectorul ajustat $\hat{\mathbf{b}} = \mathbf{P}\mathbf{b}$.

(c) Verificați că $\mathbf{P} = \mathbf{P}^T$ și $\mathbf{P}^2 = \mathbf{P}$.

Soluție. (a) Pornim de la formula soluției $\mathbf{x}^\star = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$. Apoi avem:

$$\begin{aligned} \mathbf{r} &= \mathbf{b} - \mathbf{A}\mathbf{x}^\star \\ &= \mathbf{b} - \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} \end{aligned} \tag{29}$$

Ca să verificăm dacă reziduul este orthogonal pe coloanele matrice vom calcula:

$$\mathbf{A}^T\mathbf{r} = \mathbf{A}^T\mathbf{b} - \mathbf{A}^T\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} = \mathbf{A}^T\mathbf{b} - \mathbf{A}^T\mathbf{b} = \mathbf{0}_{n\times 1}. \tag{30}$$

(b) Vectorul $\hat{\mathbf{b}} = \mathbf{A}\mathbf{x}^\star = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$ este vectorul din spațiul matrice $\mathbf{A}$ ca este cel mai aproape de vectorul $\mathbf{b}$. Deci $\mathbf{P} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ este matricea de proiecție pe spațiul coloanelor matricei $\mathbf{A}$.

(c) Din construcție $\mathbf{P}$ este simetric și idempotent; se poate verifica direct că $\mathbf{P}^T = \mathbf{P}$ și $\mathbf{P}^2 = \mathbf{P}$.

Problema 4

Cu aceleași $\mathbf{A}$ și $\mathbf{b}$ din Problema 2, calculați soluția ridge se notează

$$\mathbf{x}_\lambda = \arg\min_{\mathbf{x}} \ \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \lambda\|\mathbf{x}\|_2^2. \tag{31}$$

Găsiți $\mathbf{x}_\lambda$, ca formulă, și comparați cu soluția clasică la problema celor mai mici pătrate (soluția ne-regularizată).

Soluție. Ecuațiile normale ridge:

$$(\mathbf{A}^T\mathbf{A} + \lambda\mathbf{I})\mathbf{x}_\lambda = \mathbf{A}^T\mathbf{b}. \tag{32}$$

Acum $\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $\mathbf{A}^T\mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$$\mathbf{A}^T\mathbf{A} + \lambda\mathbf{I} = \begin{bmatrix} 2+\lambda & 1 \\ 1 & 2+\lambda \end{bmatrix}, \left(\mathbf{A}^T\mathbf{A} + \lambda\mathbf{I}\right)^{-1} = \frac{1}{(2+\lambda)^2 - 1}\begin{bmatrix} 2+\lambda & -1 \\ -1 & 2+\lambda \end{bmatrix}. \tag{33}$$

Prin urmare:

$$\mathbf{x}_\lambda^\star = \frac{1+\lambda}{(2+\lambda)^2 - 1}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\lambda+3}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \tag{34}$$

pentru că $(2+\lambda)^2 - 1 = (\lambda+1)(\lambda+3)$.

Soluția celor mai mică pătrate standard este $\mathbf{x}_0 = \frac{1}{3}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ (obținută la $\lambda = 0$).

Predicția $\hat{\mathbf{b}}_\lambda = \mathbf{A}\mathbf{x}_\lambda^\star = \begin{bmatrix} \frac{2}{\lambda+3} \\ \frac{1}{\lambda+3} \\ \frac{1}{\lambda+3} \end{bmatrix}$. Regularizarea micșorează continuu norma lui $\mathbf{x}$ pe măsură ce $\lambda \to \infty$.

Problema 5

Luând în considerare modelul regresiei liniare simple unde $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ cu $\varepsilon \in \mathcal{N}(0, \sigma^2)$, calculați regresia liniară pentru setul de date de la punctul 1. Dar pentru

$$\{(1,5),\ (2,7),\ (3,9),\ (4,11)\}. \tag{35}$$

.

Soluție. Întâi să analizăm cazul general în care avem un set de date de dimensiune N:

$$\{(x_1, y_1),\ (x_2, y_2), \ldots,\ (x_N, y_N)\}. \tag{36}$$

. Forma vectorială este

$$\mathbf{Y} = \beta_0 + \beta_1 \mathbf{X} + \varepsilon \tag{37}$$

unde zgomotul ε este independent de variabila $\mathbf{X}$. Calculăm expectanța

$$E[\mathbf{Y}] = \beta_0 + \beta_1 E[\mathbf{X}] + E[\varepsilon] = \beta_0 + \beta_1 E[\mathbf{X}] \tag{38}$$

și covarianța

$$Cov(\mathbf{X}, \mathbf{Y}) = Cov(\mathbf{X}, \beta_0 + \beta_1 \mathbf{X} + \varepsilon) = \beta_0 Cov(\mathbf{X}, 1) + \beta_1 Cov(\mathbf{X}, \mathbf{X}) + \underbrace{Cov(\mathbf{X}, \varepsilon)}_{=0,\ \text{indep.}} \tag{39}$$

$$= \sum_i (x_i - E[\mathbf{X}]) \underbrace{(1 - E[1])}_{=0,\ \text{scalar}}^\top + \beta_1 Var(\mathbf{X}) = \beta_1 Var(\mathbf{X}) \tag{40}$$

Rezolvăm sistemul prin substituție și obținem

$$\beta_1 = \frac{Cov(\mathbf{X}, \mathbf{Y})}{Var(\mathbf{X})}, \quad \beta_0 = E[\mathbf{Y}] - \beta_1 E[\mathbf{X}] \tag{41}$$

Întorcându-ne la problema noastră putem calcula rapid expectanța, covarianța și varianța:

$$\mu_x = E[\mathbf{X}] = \frac{1}{n} \sum_{i=1}^{N} x_i \tag{42}$$

$$\mu_y = E[\mathbf{Y}] = \frac{1}{n} \sum_{i=1}^{N} y_i \tag{43}$$

$$s_{xx} = Var(\mathbf{X}, \mathbf{X}) = \sum_{i=1}^{N} (x_i - E[\mathbf{X}])^2 \tag{44}$$

$$s_{xy} = Cov(\mathbf{X}, \mathbf{X}) = \sum_{i=1}^{N} (x_i - E[\mathbf{X}])(y_i - E[\mathbf{Y}]) \tag{45}$$

$$\tag{46}$$

Iar predictorii sunt calculați drept

$$\hat{\beta}_1 = \frac{s_{xy}}{s_{xx}}, \qquad \hat{\beta}_0 = \mu_y - \hat{\beta}_1 x \tag{47}$$

de unde calculăm modelul regresiei liniare $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ astfel încât atunci când primim un punct de test z răspunsul modelului devine $\hat{y}_z = \hat{\beta}_0 + \hat{\beta}_1 z$.