به نام خدا

پاسخ تمرین سری سوم

۱-

(b) i.

$$
\begin{aligned}
A &= \begin{pmatrix} 2 & 4 & 6 \\ 4 & 11 & 15 \\ 6 & 15 & 23 \end{pmatrix} \\
&= \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 2 & 4 & 6 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 3 \\ 0 & 3 & 5 \end{pmatrix} \\
&= \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 2 & 4 & 6 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 3 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} \\
&= \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 2 & 4 & 6 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 3 & 3 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 2 \end{pmatrix} \\
&= \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & 1 \end{pmatrix}}_{L} \underbrace{\begin{pmatrix} 2 & 4 & 6 \\ 0 & 3 & 3 \\ 0 & 0 & 2 \end{pmatrix}}_{U}
\end{aligned}
$$

ii. To find $A = LDL^T$, we will try to write $U = DL^T$ by pulling out a row scaling matrix D

from U to leave something unit upper triangular:

$$\begin{aligned} A &= \begin{pmatrix} 2 & 4 & 6 \\ 4 & 11 & 15 \\ 6 & 15 & 23 \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & 1 \end{pmatrix}}_{L} \underbrace{\begin{pmatrix} 2 & 4 & 6 \\ 0 & 3 & 3 \\ 0 & 0 & 2 \end{pmatrix}}_{U} \\ &= \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & 1 \end{pmatrix}}_{L} \underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}}_{D} \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}}_{L^T} \end{aligned}$$

iii. To get the Cholesky factorization of A from the factorization $A = LDL^T$, we do the following:

$$\begin{aligned} A &= LDL^T \\ &= LD^{\frac{1}{2}}D^{\frac{1}{2}}L^T \\ &= (LD^{\frac{1}{2}})(D^{\frac{1}{2}}L^T) \\ &= R^T R. \end{aligned}$$

For the matrix above, this process gives

$$\begin{aligned} A &= \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} \sqrt{2} & 0 & 0 \\ 2\sqrt{2} & \sqrt{3} & 0 \\ 3\sqrt{2} & \sqrt{3} & \sqrt{2} \end{pmatrix}}_{R^T} \underbrace{\begin{pmatrix} \sqrt{2} & 2\sqrt{2} & 3\sqrt{2} \\ 0 & \sqrt{3} & \sqrt{3} \\ 0 & 0 & \sqrt{2} \end{pmatrix}}_{R} \end{aligned}$$

Answer

Using material from the worked example in the notes we set

$$\begin{bmatrix} 3 & 1 & 6 \\ -6 & 0 & -16 \\ 0 & 8 & -17 \end{bmatrix} = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ L_{21}U_{11} & L_{21}U_{12}+U_{22} & L_{21}U_{13}+U_{23} \\ L_{31}U_{11} & L_{31}U_{12}+L_{32}U_{22} & L_{31}U_{13}+L_{32}U_{23}+U_{33} \end{bmatrix}$$

and comparing elements row by row we see that

$$\begin{array}{lll} U_{11}=3, & U_{12}=1, & U_{13}=6, \\ L_{21}=-2, & U_{22}=2, & U_{23}=-4 \\ L_{31}=0 & L_{32}=4 & U_{33}=-1 \end{array}$$

and it follows that

$$\begin{bmatrix} 3 & 1 & 6 \\ -6 & 0 & -16 \\ 0 & 8 & -17 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 6 \\ 0 & 2 & -4 \\ 0 & 0 & -1 \end{bmatrix}$$

is an LU decomposition of the given matrix.

We found earlier that the coefficient matrix is equal to $LU = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 6 \\ 0 & 2 & -4 \\ 0 & 0 & -1 \end{bmatrix}$.

First we solve $LY = B$ for Y, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 17 \end{bmatrix}.$$

The top line implies that $y_1 = 0$. The middle line states that $-2y_1 + y_2 = 4$ and therefore $y_2 = 4$. The last line tells us that $4y_2 + y_3 = 17$ and therefore $y_3 = 1$.

Finally we solve $UX = Y$ for X, we have

$$\begin{bmatrix} 3 & 1 & 6 \\ 0 & 2 & -4 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 1 \end{bmatrix}.$$

The bottom line shows that $x_3 = -1$. The middle line then shows that $x_2 = 0$, and then the top line gives us that $x_1 = 2$. The required solution is $X = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$.

$$
\begin{aligned}
\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} &= \begin{bmatrix} -1 & -1 & 1 \\ 1 & 3 & 3 \\ -1 & -1 & 5 \\ 1 & 3 & 7 \end{bmatrix} \\
&= \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ 0 & R_{22} & R_{23} \\ 0 & 0 & R_{33} \end{bmatrix}
\end{aligned}
$$

First column of Q and R

$$
\tilde{q}_1 = a_1 = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \qquad R_{11} = \|\tilde{q}_1\| = 2, \qquad q_1 = \frac{1}{R_{11}}\tilde{q}_1 = \begin{bmatrix} -1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix}
$$

Second column of Q and R

- compute $R_{12} = q_1^T a_2 = 4$
- compute

$$
\tilde{q}_2 = a_2 - R_{12}q_1 = \begin{bmatrix} -1 \\ 3 \\ -1 \\ 3 \end{bmatrix} - 4 \begin{bmatrix} -1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}
$$

- normalize to get

$$
R_{22} = \|\tilde{q}_2\| = 2, \qquad q_2 = \frac{1}{R_{22}}\tilde{q}_2 = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}
$$

Third column of Q and R

- compute $R_{13} = q_1^T a_3 = 2$ and $R_{23} = q_2^T a_3 = 8$
- compute

$$\tilde{q}_3 = a_3 - R_{13}q_1 - R_{23}q_2 = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \end{bmatrix} - 2 \begin{bmatrix} -1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix} - 8 \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \\ 2 \\ 2 \end{bmatrix}$$

- normalize to get

$$R_{33} = \|\tilde{q}_3\| = 4, \qquad q_3 = \frac{1}{R_{33}}\tilde{q}_3 = \begin{bmatrix} -1/2 \\ -1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$$

Final result

$$\begin{aligned} \begin{bmatrix} -1 & -1 & 1 \\ 1 & 3 & 3 \\ -1 & -1 & 5 \\ 1 & 3 & 7 \end{bmatrix} &= \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ 0 & R_{22} & R_{23} \\ 0 & 0 & R_{33} \end{bmatrix} \\ &= \begin{bmatrix} -1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 \\ -1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & 4 & 2 \\ 0 & 2 & 8 \\ 0 & 0 & 4 \end{bmatrix} \end{aligned}$$

Application to linear regression

The QR method is often used to estimate linear regressions.

In a linear regression we have an $N \times 1$ vector y of outputs and an $N \times K$ matrix of inputs whose columns are assumed to be linearly independent. We need to find the $K \times 1$ coefficient vector β that minimizes the mean squared errors made by using the fitted values

$$\widehat{y} = X\beta$$

to predict the actual values y.

The well-known solution to this problem is the so-called ordinary least squares (OLS) estimator

$$\beta = (X^\intercal X)^{-1} X^\intercal y$$

We can simplify the formula for the OLS estimator, avoid to invert a matrix and thus reduce the computational burden (and the possible numerical instabilities) by computing the QR decomposition of X:

$$X = QR$$

where Q is $N \times K$ and R is $K \times K$.

Then, the OLS estimator becomes

$$\begin{aligned}\beta &= (X^\intercal X)^{-1} X^\intercal y \\ &= (R^\intercal Q^\intercal Q R)^{-1} R^\intercal Q^\intercal y \\ &= (R^\intercal R)^{-1} R^\intercal Q^\intercal y \\ &= R^{-1} (R^\intercal)^{-1} R^\intercal Q^\intercal y \\ &= R^{-1} Q^\intercal y\end{aligned}$$

or

$$R\beta = Q^\intercal y$$

The latter way of writing the solution is more convenient: since R is upper triangular, we do not need to invert it, but we can use the back-substitution algorithm to find the solution β.

-۵

$A = A_p \cdots A_2 A_1$ A_iها مربعی‌اند

$Q_0 = I$

$\tilde{A}_1 \triangleq A_1 Q_0 \xrightarrow{\text{تجزیه } QR} \tilde{A}_1 = Q_1 R_1 \longrightarrow Q_1^T \tilde{A}_1 = R_1 \longrightarrow Q_1^T A_1 Q_0 = R_1$ بالامثلثی

$\tilde{A}_2 \triangleq A_2 Q_1 \xrightarrow{\text{تجزیه } QR} \tilde{A}_2 = Q_2 R_2 \longrightarrow Q_2^T (A_2 Q_1) = R_2$ بالامثلثی

$\vdots$

$\tilde{A}_p \triangleq A_p Q_{p-1} \xrightarrow{\text{تجزیه } QR} \tilde{A}_p = Q_p R_p \longrightarrow Q_p^T (A_{p-1} Q_{p-1}) = R_p$ بالامثلثی

$$A = A_p A_{p-1} \cdots A_2 A_1 = A_p \underbrace{Q_{p-1} Q_{p-1}^T}_{I} \underbrace{A_{p-1} Q_{p-2}}_{R_{p-1}} \cdots \underbrace{Q_2 Q_2^T}_{I} A_2 \underbrace{Q_1 Q_1^T}_{I} \underbrace{A_1 Q_0}_{R_1} \quad (Q_0 = I)$$

$$= A_p Q_{p-1} R_{p-1} \cdots R_2 R_1$$

$$\longrightarrow Q_p^T A = Q_p^T A_p Q_{p-1} R_{p-1} \cdots R_1 = R_p \cdots R_1 \longrightarrow A = Q_p (R_p \cdots R_1) \xRightarrow{A = QR} \begin{cases} Q = Q_p \\ R = R_p \cdots R_1 \end{cases}$$

-٦

QR تجزیه $\rightarrow A = QR \Rightarrow |\det(A)| = |\det(Q)|\,|\det(R)| = |r_1(1)| \cdots |r_n(n)| \overset{(*)}{\leq} \|a_1\|_2 \cdots \|a_n\|_2$

$R = [r_1, \dots, r_n]$

(*): نشان می‌دهیم برای $i \in \{1, \dots, n\}$ داریم: $|r_i(i)| \leq \|a_i\|_2$

$Q^T A = R \Rightarrow Q^T a_i = r_i \Rightarrow \|Q^T a_i\|_2 = \|r_i\|_2 \Rightarrow \|a_i\|_2 = \|r_i\|_2 \geq |r_i(i)|$ ☑

اثبات این تجزیه به راحتی از رابطه ضرب ماتریس بلوکی روشن است.

الف – ۲) با تعریف دترمینان نشان دهید:

$$\det \underbrace{\begin{bmatrix} I_{m_1} & & & & \\ & \ddots & & & \\ & & A_j & & \\ & & & \ddots & \\ & & & & I_{m_k} \end{bmatrix}}_{S_j} = \det(A_j)$$

اثبات به کمک رابطه بازگشتی دترمینان روشن است. (به جای اندیس j از اندیس i نوشته شده).

بسط نسبت به سطر ۱ — تعداد زیرماتریس ها

$$\det(S_i) = (-1)^{1+1} \times 1 \times \det(S_i(2{:}end, 2{:}end)) = \det(S_i(3{:}end, 3{:}end)) = \dots = \det(S_i(m_1+\dots+m_{i-1}+1{:}end, m_1+\dots+m_{i-1}+1{:}end))$$

$$= \det(S_i(m_1+\dots+m_{i-1}+1{:}end-1, m_1+\dots+m_{i-1}+1{:}end-1)) = \dots =$$

از این جا به بعد از سطر آخر بسط داده و بالا آمده

$$= \det(S_i(m_1+\dots+m_{i-1}+1{:}end-m_n-\dots-m_{i+1}, m_1+\dots+m_{i-1}+1{:}end-m_n-\dots-m_{i+1}))$$

$$= \det(S_i(\underset{A_i}{\text{شماره سطرهای}}, \underset{A_i}{\text{شماره ستون‌های}})) = \det(A_i)$$

الف – ۳) نشان دهید:

$$\det(A) = \det(A_1)\det(A_2)\dots\det(A_k)$$

طبق قسمت قبل $\longrightarrow$ $\det(A) = \det(S_1)\dots\det(S_n) = \det(A_1)\dots\det(A_n)$

ب) فرض کنید $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ که $A \in \mathbb{R}^{p\times p}, B \in \mathbb{R}^{p\times q}, C \in \mathbb{R}^{q\times p}, D \in \mathbb{R}^{q\times q}$

ب – ۱) تجزیه‌های شبه LDU زیر را برای ماتریس بلوکی M ثابت کنید. یعنی اگر A معکوس‌پذیر باشد، نشان دهید:

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ CA^{-1} & I_q \end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{bmatrix}\begin{bmatrix} I_p & A^{-1}B \\ 0 & I_q \end{bmatrix}$$

و اگر D معکوس‌پذیر باشد، نشان دهید:

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I_p & BD^{-1} \\ 0 & I_q \end{bmatrix}\begin{bmatrix} A - BD^{-1}C & 0 \\ 0 & D \end{bmatrix}\begin{bmatrix} I_p & 0 \\ D^{-1}C & I_q \end{bmatrix}$$

ب – ۲) نشان دهید اگر A و D معکوس‌پذیر باشند داریم:

$$\det\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A)\det(D - CA^{-1}B) = \det(D)\det(A - BD^{-1}C)$$

ب – ۱) به راحتی هر دو قسمت از ضرب بلوکی ماتریس ها روشن است.

ب-۲) از طرفین روابط قسمت (ب-۱) دترمینان می‌گیریم. از آن جا که دترمینان ماتریس‌های مثلثی حاصل ضرب عناصر قطر است و دترمینان ماتریس بلوکی قطری برابر حاصل ضرب دترمینان بلوک‌های روی قطر آن است، هر دو رابطه مستقیم به دست می‌آیند.

ج) فرض کنید $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times m}, X \in \mathbb{R}^{m\times m}$ که X معکوس‌پذیر است.

ج – ۱) قضیه دترمینان Sylvester را اثبات کنید:

$$\det(I_m + AB) = \det(I_n + BA)$$

ج – ۲) فرم کلی‌تر قضیه‌ی بالا را اثبات کنید:

$$\det(X + AB) = \det(X)\det(I_n + BX^{-1}A)$$

ج – ۳) برای دو بردار $u, v \in \mathbb{R}^m$ با استفاده از قضیه‌ی بالا نشان دهید:

$$\det(X + uv^T) = (1 + v^T X^{-1} u)\det(X)$$

ج-۱)

$$M = \begin{pmatrix} I_m & -A \\ B & I_n \end{pmatrix} \Rightarrow \det(M) \overset{\text{از قسمت قبل}}{=\!=} \det(I_m)\det(I_n + B I_m^{-1} A) = \det(I_n)\det(I_m + A I_n^{-1} B) \Rightarrow \square$$

ج-۲)

$$\det(X + AB) = \det(X(I_m + X^{-1}AB)) = \det(X)\det(I_m + \underbrace{X^{-1}}A B) \overset{(ج-۱)}{=} \det(X)(I_n + BX^{-1}A) \quad \square$$

ج-۳)

$(A \triangleq u \quad B \triangleq v^T) \rightarrow$ جایگذاری در ج-۲