97

Engineering Mathematics Basics
Matrix Basics

By *Professor Qingjiang Shi*

Zhicheng Liu

February 13, 2021

1. Let A be the $n \times n$ matrix with $a_{j,j+1} = 1, j = 1, ..., n-1$ and all other elements zero. Represent A as a sum of vector outer products.

   solution

   $A = \sum_{j=1}^{j=n-1} I_j I_j$

2. Consider the block matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ with $A \in C^{n\times n}$ nonsingular.

   1) Determine $X \in C^{m\times m}$ and $Y \in C^{n\times n}$ such that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & O \\ CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ O & X \end{bmatrix} (blockLUfactorization)$$

   and that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & BD^{-1} \\ O & I \end{bmatrix} \begin{bmatrix} Y & O \\ C & D \end{bmatrix} (blockLUfactorization)$$

   solution:

   prove the first equation:if $X = D - CA^{-1}B$,

$$\begin{bmatrix} I & O \\ CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ O & X \end{bmatrix} = \begin{bmatrix} IA & IB \\ CA^{-1}A & CA^{-1}B + IX \end{bmatrix} = \begin{bmatrix} A & B \\ C & CA^{-1}B + X \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

   prove the second equation:if $Y = A - BD^{-1}C$,

$$\begin{bmatrix} I & BD^{-1} \\ O & I \end{bmatrix} \begin{bmatrix} Y & O \\ C & D \end{bmatrix} = \begin{bmatrix} IY + BD^{-1}C & BD^{-1}D \\ IC & ID \end{bmatrix} = \begin{bmatrix} Y + BD^{-1}C & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

   2) Prove that

$$det(A)det(D - CA^{-1}B) = det(D)det(A - BD^{-1}C) \tag{1}$$
$$det(I - CB) = det(I - BC) \tag{2}$$
$$det(I - xy^*) = 1 - y^*x \tag{3}$$

   where $x, y \in C^n$.

   solution:

   prove the first equation:

$$\begin{bmatrix} I & O \\ CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ O & X \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & BD^{-1} \\ O & I \end{bmatrix} \begin{bmatrix} Y & O \\ C & D \end{bmatrix}$$

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

$$det(I)det(AX) = det(I)det(YD)$$

$$det(AX) = det(YD)$$

   ,$X = D - CA^{-1}B$,$Y = A - BD^{-1}C$

$$det(A)det(D - CA^{-1}B) = det(D)det(A - BD^{-1}C)$$

prove the second equation:

if A=D=I,

$$det(I)det(I-CI^{-1}B)=det(I)det(I-BI^{-1}C)$$

$$det(I-CB)=det(I)det(I-BC)$$

prove the last equation:

if C=x,B=y^*,

$det(I-xy^*)=det(I)det(I-y^*x)$

$det(I-xy^*)=det(I-y^*x)$

Because A is a 1×1 matrix,

$det(I-xy^*)=1-y^*x$

3. Let $A\in C^{n\times n}$, the trace of A is defined as the sum of its diagonal elements, i.e., $trace(A)=\sum_{i=1}^{n}a_{ii}$.

   1) Show that the trace is a linear function, i.e., if $A,B\in C^{n\times n}$ and $\alpha,\beta\in C$, then

$$trace(\alpha A+\beta B)=\alpha trace(A)+\beta trace(B).$$

   solution:

   $trace(\alpha A+\beta B)=\sum_{i=1}^{n}(\alpha a_{ii}+\beta b_{ii})=\alpha\sum_{i=1}^{n}a_{ii}+\beta\sum_{i=1}^{n}b_{ii}=\alpha trace(A)+\beta trace(B)$

   2) Show that $trace(AB)=trace(BA)$, even though in general $AB\neq BA$.

   solution:

   $trace(AB)=\sum_{i=1}^{n}\sum_{k=1}^{n}a_{ik}b_{ki})$

   $trace(BA)=\sum_{i=1}^{n}\sum_{k=1}^{n}b_{ik}a_{ki}=\sum_{k0=1}^{n}\sum_{i0=1}^{n}b_{k0i0}a_{i0k0}=\sum_{k0=1}^{n}\sum_{i0=1}^{n}a_{i0k0}b_{k0i0}=trace(AB)$

   3) Show that if $S\in R^{n\times n}$ is skew-symmetric, i.e., $S^T=-S$, the $trace(S)=0$. Prove the converse to this statement or provide a counterexample.

   solution:

   $trace(S)=trace(S^T)=trace(-S)=-trace(S)$

   So,trace(S)=0.

   counterexample of the converse to this statement:

   $S=\begin{bmatrix}0&1\\0&0\end{bmatrix},trace(\begin{bmatrix}0&1\\0&0\end{bmatrix})=0$

   $S^T=(\begin{bmatrix}0&1\\0&0\end{bmatrix})^T=(\begin{bmatrix}0&0\\1&0\end{bmatrix})\neq(\begin{bmatrix}0&-1\\0&0\end{bmatrix})=-S$

4. Let $\lambda_1,\lambda_2,...,\lambda_n$ be the eigenvalues of A. Show that

$$det(A)=\lambda_1\lambda_2...\lambda_n, \tag{1}$$
$$trace(A)=\lambda_1+\lambda_2+...+\lambda_n, \tag{2}$$
$$trace(A^k)=\lambda_1^k+\lambda_2^k+...+\lambda_n^k, k=1,2,... \tag{3}$$

   solution:

prove the first equation: $A = C^{-1} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} C$

$det(A) = det(C^{-1})det(\begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix})det(C)$

$det(A) = det(C^{-1}C)det(\begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix})$

$det(A) = \lambda_1 \cdots \lambda_n$

prove the second equation:

$det(\lambda - A) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n)$

the coefficient of the term of λ^{n-1} is $-(\lambda_1 + \cdots + \lambda_n)$

$det(\lambda - A) = det(\begin{bmatrix} \lambda - a_{1,1} & \cdots & -a_{1,n} \\ \vdots & \ddots & \vdots \\ -a_{n,1} & \cdots & \lambda - a_{n,n} \end{bmatrix})$

the coefficient of the term of λ^{n-1} is only related to $(\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn})$

so,the coefficient of the term of λ^{n-1} is $-(a_1 + \cdots + a_n)$

so,$-(\lambda_1 + \cdots + \lambda_n) = -(a_1 + \cdots + a_n)$

prove the last equation:

if $A\alpha = \lambda_1 \alpha, A^k = \lambda^k \alpha$

$So, \lambda_1^k, \lambda_2^k, ..., \lambda_n^k$ be the eigenvalues of A^k.

$So, trace(A^k) = \lambda_1^k + \lambda_2^k + ... + \lambda_n^k, k = 1, 2, ...$

5. 1) Let $x_1, x_2, ..., x_k$ be eigenvectors of A. Show that $S = span\{x_1, x_2, ..., x_k\}$ is invariant under A.

solution:

if $y \in S$,

$\exists a_1, \cdots a_k, y = a_1 x_1 + \cdots a_k x_k$

$so, Ay = a_1 \lambda_1 x_1 + \cdots + a_k \lambda_k x_k$

so,S is invariant under A.

2) Let $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Show that the space $S = span\{e_1, e_2\}$ is invariant under A and is not spanned by eigenvectors of A. Here $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

solution:

$\forall y \in S$, y=$c_1e_1 + c_2e_2$

$Ay = Ac_1e_1 + Ac_2e_2 = c_1Ae_1 + c_2Ae_2 = c_1 \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} + c_1 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = (2c_1 + c_2)e_1 + 2c_2e_2$

so,S is invariant under A.

$det(\lambda E - A) = 0$ can derive $\lambda_1 = \lambda_2 = 2, \lambda_3 = 1$

$(2E - A)\alpha = 0$ can derive $\alpha_1 = k_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

$(E - A)\alpha = 0$ can derive $\alpha_2 = k_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

$e_2 = b_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ can derive that b_1, b_2 has no solution.

so,S is not spanned by eigenvectors of A.

6. Verify that $\| xy^* \|_F = \| xy^* \|_2 = \| x \|_2 \| y \|_2$ for any $x, y \in C^n$.

   solution:

   $\| xy^* \|_F^2 = trace((yx^*)(xy^*)) = (x^*x)trace(yy^*) = \sum_{i=1}^n x_i^2 \sum_{j=1}^n y_j^2$

   $\| xy^* \|_2 = \max_z((\| xy^*z \|_2^2)/(\| z \|_2^2)) = \max_z((|y^*z| \| x \|_2^2)/(\| z \|_2^2)) = \| y \|_2^2 \| ky \|_2^2 \| x \|_2^2 \ / \ \| ky \|_2^2 = \| x \|_2^2 \| y \|_2^2$,k is a number.

7. Let $A \in C^{n\times m}$. Show that $\| A \|_2 \leq \| A \|_F$. (Recall that $trace(B) = \sum_i \lambda_i(B)$).

   solution:

   $\| A \|_F^2 = trace(A^*A)$

   $\| A \|_2^2 = \max_x \frac{\|Ax\|_2^2}{\|x\|_2^2} = \max_i \lambda_i$

   $trace(A^*A) = \sum_{i=1}^n \lambda_i > \max_i \lambda_i$

8. For $A \in C^{n\times m}$, show that

$$\| A \|_1 = \underset{x\neq 0}{max} \frac{\| Ax \|_1}{\| x \|_1} = \underset{1\leq j\leq n}{max} \sum_{i=1}^m | a_{ij} | \tag{1}$$

$$\| A \|_\infty = \underset{x\neq 0}{max} \frac{\| Ax \|_\infty}{\| x \|_\infty} = \underset{1\leq j\leq n}{max} \sum_{i=1}^m | a_{ij} | \tag{2}$$

   solution:

   $\| A \|_1 = \max_{x\neq 0} \frac{\|Ax\|_1}{\|x\|_1} = \max_{x\neq 0} \frac{\sum_i |\sum_j a_{ij}x_i|}{\sum_i |x_i|} \leq \max_{x\neq 0} \frac{\sum_i \sum_j |a_{ij}x_i|}{\sum_i |x_i|} \leq \max_{x\neq 0} \frac{(\max_j \sum_i (a_{ij})) \sum_j |x_i|}{\sum_i |x_i|} = \max_j \sum_i (a_{ij})$

   The proof of the second equation is similar.

9. Let $A, B \in R^{m\times n}$. Recall that $\| A \|_F^2 = trace(A^TA)$ and that for any matrix D for which the product AD is defined, $trace(AD) = trace(DA)$.

   1) Show that the Q that minimizes $\| A - QB \|_F$ over all choices of orthogonal Q also maximizes $trace(A^TQB)$.

solution:

$\| A-QB \|_F^2 = trace((A-BQ)^T(A-QB)) = trace((A^T-B^TQ^T)(A-QB)) = trace(A^TA+B^TQ^TQB-A^TQB-B^TQ^TA) = trace(A^TA+B^TB)-2trace(A^TQB)$

$trace(A^TA+B^TB)$ is constant.

so,the Q that minimizes $\| A-QB \|_F$ over all choices of orthogonal Q also maximizes $trace(A^TQB)$

2) Suppose that the SVD of the $m \times m$ matrix BA^T is $U\Sigma V^T$, where U and V are $m \times m$ and orthogonal and Σ is diagonal with diagonal entries $\sigma_1 \geq ... \geq \sigma_m \geq 0$. Define $Z = V^TQU$. Use these definitions and (i) to show that

$$trace(A^TQB) = trace(Z\Sigma) \leq \sum_{i=1}^{m} \sigma_i$$

.

solution:

$trace(A^TQB) = trace(QBA^T) = trace(QU\Sigma V^T) = trace(V^TQU\Sigma) = trace(Z\Sigma)$

Z is orthogonal and diagonal,hence z_{ij}=0 or 1 or -1, $trace(Z\Sigma) = \sum_i z_{ii}\sigma_i \leq \sum_i \sigma_i$

3) Identify the choice of Q that gives equality in the bound of (ii).

solution:

if $Q = VU^T$,

$Z = V^TQU = V^TVU^TU = V^{-1}VU^{-1}U = II = I$

$trace(Z\Sigma) = trace(\Sigma) = \sum_i \sigma_i$

4) Carefully state a theorem summarizing the solution to the minimization of $\| A-QB \|_F$.

solution:

$Q = VU^T, Q^T = Q^{-1}, \| A-QB \|_F$ is minimized by Q $\Rightarrow BA^T = U\Sigma V^T$is an SVD.

10. Let $A \in C^{m\times n}$.

1) Show that $(range(A))^{\perp} = null(A^*)$ and that $(null(A))^{\perp} = range(A^*)$.

solution:

$A = U\Sigma V^*, r = Rank(A) \Rightarrow null(A^*)$=$span\{u_{r+1}, \cdots, u_m\}, range(A) = span\{u_1, \cdots, u_r\} \Rightarrow (range(A))^{\perp} = null(A) \Rightarrow range(A) = (null(A))^{\perp}$

2) Show respectively that $AA^+, A^+A, I-A^+AandI-AA^+$ are the orthogonal projectors onto $range(A)$, $range(A^*), null(A)andnull(A^*)$.

solution:

First part:

$(AA^+)^2 = (AA^+A)A^+ = AA^+$

$\forall x \in range(A), x = Ay = AA^+Ay = AA^+x \Rightarrow x \in range(AA^+)$

$\forall x \in range(AA^+), x = AA^+y = A(A^+y) \Rightarrow x \in range(A)$

hence AA^+ are the orthogonal projectors onto $range(A)$.

Second part:

By first part,$A^*(A^*)^+$ are the orthogonal projectors onto $range(A^*)$.

hence A^+A are the orthogonal projectors onto $range(A^*)$.

Third part:

By first part,$I - AA^+$ are the orthogonal projectors onto $range(A^{\perp}) = null(A)$.

hence $I - AA^+$ are the orthogonal projectors onto $null(A)$.

Last part:

By second part,$I - A^+A$ are the orthogonal projectors onto $range((A^*)^{\perp}) = null(A^*)$.

hence $I - AA^+$ are the orthogonal projectors onto $null(A^*)$.

3) Suppose $rank(A) = r$, let $A = U\Sigma V^*$ with $U = [U_1 \quad U_2]$ and $V = [V_1 \quad V_2]$ unitary and partitioned, such that $U_1 \in C^{m\times r}$ and $V_1 \in C^{n\times r}$. Show that $U_1U_1^*, V_1V_1^*, V_2V_2^*$ and $U_2U_2^*$ are the orthogonal projectors onto $range(A), range(A^*), null(A)$ and $null(A^*)$, respectively.

solution:

$$AA^+ = U\Sigma V^*(V\Sigma^+U^*) = U \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} U^* = U1U1^*$$

hence $U1U1^*$ are the orthogonal projectors onto $range(A)$.

other parts are similar.