Matrix handbook for Statistics and Machine Learning

Thu Nguyen
University of Louisiana at Lafayette

September 25, 2019

Abstract

This handbook contains popular identities, inequalities and other matrix-related topics. Most of the proof for the statements in the book, without any further comments, can be found in [5], its solution manual or Wikipedia.

Contents

1 Notations

- $diag(a_1, ..., a_m)$: a diagonal matrix whose diagonal entries are $a_1, .., a_m$,
- $x \in \mathbb{R}^n$: a column vector x with n entries.
- $x \in \mathbb{R}^{n\times 1}$: a column vector x with n entries.

2 Trace

Let $A = (a_{ij}), B$ be $m \times m$ matrices, then trace of A is defined as $trace(A) = \sum_{i=1}^{m} a_{ii}$. We have

1. $trace(cA) = c\, trace(A)$
2. $trace(A \pm B) = trace(A) \pm trace(B)$
3. $trace(AB) = trace(BA)$
4. $trace(AA') = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}^2$

Inequality for trace from [3], chapter 10,

1. Suppose A is an $m \times m$ nonnull nonnegative definite matrix and $\frac{1}{p} + \frac{1}{q} = 1$, where $p > 1$. Then
$$trace(AB) \leq \{trace(A^p)\}^{1/p}$$
for every $m \times m$ nonnegative definite matrix B such that $trace(B^q) = 1$. Equality if and only if $B^q = \{trace(A^p)\}^{-1} A^p$.

2. Suppose A, B are $m \times m$ nonnull nonnegative definite matrices and α is a scalar such that $0 < \alpha < 1$. Then
$$trace(A^{\alpha}B^{1-\alpha}) \leq \{trace(A)\}^{\alpha}\{trace(B)\}^{1-\alpha}$$
with equality if and only if $B = cA$ for some positive scalar c.

3. Suppose A, B are $m \times m$ nonnull nonnegative definite matrices and $p > 1$. Then
$$[trace\{(A+B)^p\}]^{1/p} \leq \{trace(A^p)\}^{1/p} + \{trace(B^p)\}^{1/p}$$
with equality if and only if $B = cA$ for some positive scalar c.

4. Suppose A is an $m \times m$ nonnegative definite matrix. Then,
   - $\frac{1}{m}trace(A) \geq |A|^{1/m}$, with equality if and only if $A = cI_m$ for some $c > 0$.
   - $m^{-1}trace(AB) \geq |A|^{1/m}$ for every positive definite $m \times m$ matrix B such that $|B| = 1$, with equality if and only if A is nonsingular and $B = |A|^{1/m}A^{-1}$.

5. Let A be an $m \times m$ matrix with real eigenvalues. Then
   - $\{trace(A)\}^2 \leq m\ trace(A^2)$ with equality if and only if the eigenvalues are all equal.
   - $trace(A^2) \leq trace(A'A)$ with equality if and only if A is symmetric.

6. If A, B are $m \times n$ matrices, then
$$trace\{(A'B)^2\} \leq trace\{(A'A)(B'B)\},$$
with equality if and only if AB' is symmetric.

7. If A, B are $m \times m$ positive definite matrices, and $0 < \alpha < 1$ then
$$trace[\{\alpha A + (1-\alpha)B\}^{-1}] \leq \alpha\ trace(A^{-1}) + (1-\alpha)trace(B^{-1}).$$

3 Determinant

Determinant equality:

For a $m \times m$ real matrix A,

- $|cA| = c^n|A|$
- $|A^T| = |A|$
- $|AB| = |A|\ |B|$
- $|I + uv^T| = 1 + u^T v$ for vectors u, v of revelant sizes.
- $|I + \epsilon A| \approx 1 + |A| + \epsilon\ trace(A) + \frac{1}{2}\epsilon^2 trace(A)^2 - \frac{1}{2}\epsilon^2 trace(A^2)$ for small ϵ

1. $|I_m + AB| = |I_n + BA|$ where A, B are $m \times n, n \times m$ matrices, respectively. Therefore, the nonzero eigenvalues of AB are the same as the nonzero eigenvalues of BA.
2. $|A + cd'| = |A|(1 + d'A^{-1}c)$ where $A \in \mathbb{R}^{m \times m}$ is nonsinglar and $c, d \in \mathbb{R}^m$.

Inequalities for the determinant

1. $|A+B| \geq |A|+|B|$ where A, B are $m \times m$ nonnegative definite matrices. The equality happens if and only if $A = 0$ or $B = 0$ or $A + B$ is singular.

   Moreover, if B is also positive definite and $A - B$ is nonnegative definite, then $|A| \geq |B|$ with equality if and only if $A = B$.
2. If A is an $m \times m$ positive definite matrix, then

$$|A| \leq \prod_{i=1}^{m} A_{ii}$$

   with equality if and only if A is a diagonal matrix,
3. If B is an $m \times m$ nongsingular matrix then

$$|B|^2 \leq \prod_{i=1}^{m} \left(\sum_{j=1}^{m} b_{ij}^2 \right),$$

   with equality if and only if the rows of B are orthogonal. This can also be expressed as

$$|A| \left(\prod_{i=1}^{m} 1 \right) \leq |A \odot I_m|,$$

4. Let A, B be $m \times m$ nonnegative definite matrices, then

$$|A| \prod_{i=1}^{m} b_{ii} \leq |A \odot B|$$

5. Let A be an $m \times m$ positive definite matrix and B be an $m \times m$ nonnegative definite matrix. Then,
   - $|A + B| \geq |A|$ with equality if and only if B = 0,
   - If B is also positive definite and A B is nonnegative definite, then $|A| \geq |B|$ with equality if and only if $A = B$.

Results from [3], chapter 10,

1. Suppose that both A and B are $m \times n$ matrices. Then,

$$|A'B|^2 \leq |A'A||B'B|,$$

with equality if and only if $rank(A) < n$ or $rank(B) < n$, or $B = AC$ for some nonsingular matrix C.

2. Suppose that both A and B are $m \times m$ nonnegative definite matrices and α is a scalar such that $0 < \alpha < 1$. Then,
$$|A|^{\alpha}|B|^{1-\alpha} \leq |\alpha A + (1-\alpha)B|,$$
with equality if and only if $A = B$ or $\alpha A + (1-\alpha)B$ is singular.

3. Suppose that both A and B are nonnull $m \times m$ nonnegative definite matrices. Then
$$|A + B|^{1/m} \geq |A|^{1/m} + |B|^{1/m},$$
with equality if and only if $A + B$ is singular or $A = cB$ for $c > 0$.

4 Eigenvalues, eigenvectors

1. The characteristic polynomial of an n-by-n matrix A, being a polynomial of degree n, can be factored as follows
$$|A - \lambda I| = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda),$$
2. trace$(A) = \sum_{i=1}^{n} \lambda_i$,
3. det$(A) = \prod_{i=1}^{n} \lambda_i$,
4. The eigenvalues of the A^k, for any positive integer k, are $\lambda_1^k, ..., \lambda_n^k$.
5. If A is invertible, then the eigenvalues of A^{-1} are $\frac{1}{\lambda_1}, ..., \frac{1}{\lambda_n}$ and each eigenvalue's geometric multiplicity coincides. Moreover, since the characteristic polynomial of the inverse is the reciprocal polynomial of the original, the eigenvalues share the same algebraic multiplicity.
6. If A is not only Hermitian but also positive-definite, positive-semidefinite, negative-definite, or negative-semidefinite, then every eigenvalue is positive, non-negative, negative, or non-positive, respectively.
7. If A is unitary, every eigenvalue has absolute value $|\lambda_i| = 1$.
8. For a polynomial P, the eigenvalues of matrix $P(A)$ are $\{P(\lambda_1), \ldots, P(\lambda_k)\}$.
9. The eigenvalues of A^T are the same as the eigenvalues of A,
10. If A is triangular then the diagonal elements of A are the eigenvalues of A,
11. The eigenvalues of A is the same as of BAB^{-1} if B is invertible,
12. $rank(A) \geq r$ where A is an $m \times m$ matrix that has r nonzero eigenvalues.
13. Let λ be any eigenvalue of a $m \times m$ matrix A. Then $|\lambda| \leq \min\{\rho(A), \upsilon(A)\}$, where
$$\rho(A) = \max\left\{\sum_{j=1}^{n} |A|_{ij} : 1 \leq i \leq n\right\}, \qquad \upsilon(A) = \max\left\{\sum_{i=1}^{n} |A|_{ij} : 1 \leq j \leq n\right\}.$$

This result follows from Gerschgorin's Disk Theorem. (Proof: see [1], section 5.3)

5 Vectors

The followings are from [2], section 2.7,

- **Cauchy-Schwaz inequality:** Let $b, d \in \mathbb{R}^{p\times 1}$ then

$$(b'd)^2 \leq (b'b)(d'd)$$

  with equality if and only if $b = cd$ for some constant c.

- **Extended Cauchy-Schwaz inequality:** Let $b, d \in \mathbb{R}^{p\times 1}$, and let B be a positive definite matrix. Then

$$(b'd)^2 \leq (b'Bb)(d'B^{-1}d)$$

  with equality if and only if $b = cB^{-1}d$ for some constant c.

- Let $d \in \mathbb{R}^{p\times 1}$ be a given vector, and let B be a positive definite matrix. Then for an abitrary nonzero vector $x \in \mathbb{R}^{p\times 1}$,

$$\max_{x\neq 0} \frac{(x'd)^2}{x'Bx} = d'B^{-1}d$$

  with equality if and only if $b = cB^{-1}d$ for some constant $c \neq 0$.

Rayleigh quotient: For any vector $x \in \mathbb{R}^m$ Rayleigh quotient is defined by $R(x, A) = \frac{x'Ax}{x'x}$. From [1], section 6.9,

$$\min_{x\neq 0} R(x, A) = \min_i \lambda_i, \qquad \max_{x\neq 0} R(x, A) = \max_i \lambda_i$$

1. Let A, B be $m \times m$ positive definite matrices. Then, for any $y \neq 0$

$$y'(A + B)^{-1}y \leq \frac{(y'A^{-1}y)(y'B^{-1}y)}{y'(A^{-1} + B^{-1})y}$$

6 Norm

For all scalars $\alpha \in K$ and for all matrices $A, B \in K^{m\times n}$,

- $\|\alpha A\| = |\alpha|\|A\|$
- $\|A + B\| \leq \|A\| + \|B\|$
- $\|A\| \geq 0$

- $\|A\| = 0$ iff $A = 0$

Important norms:

- $\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|$, which is simply the maximum absolute column sum of A,
- $\|A\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|$, which is simply the maximum absolute row sum of A,
- $\|A\|_2 = \sigma_{\max}(A)$, where $\sigma_{\max}(A)$ is the largest singular value of A.
- When $p = q = 2$ for the $L_{p,q}$ norm also called the Frobenius norm, and we have:

$$\|A\|_{\mathrm{F}} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} = \sqrt{\mathrm{trace}\left(A^*A\right)} = \sqrt{\sum_{i=1}^{\min\{m,n\}} \sigma_i^2(A)},$$

  where $\sigma_i(A)$ are the singular values of A.

- $L_{p,q}$ norm, $p, q \geq 1$, defined by

$$\|A\|_{p,q} = \left(\sum_{j=1}^{n}\left(\sum_{i=1}^{m} |a_{ij}|^p\right)^{\frac{q}{p}}\right)^{\frac{1}{q}}.$$

Properties: For $A \in \mathbb{R}^{m\times r}, B \in \mathbb{R}^{r\times n}$,

- $\|A^*A\|_{\mathrm{F}} = \|AA^*\|_{\mathrm{F}} \leq \|A\|_{\mathrm{F}}^2$
- $\|A+B\|_{\mathrm{F}}^2 = \|A\|_{\mathrm{F}}^2 + \|B\|_{\mathrm{F}}^2 + 2\langle A, B\rangle_{\mathrm{F}}$, where $\langle A, B\rangle_{\mathrm{F}}$ is the Frobenius inner product.
- $||AB||_F \leq ||A||_2||B||_F \leq ||A||_F||B||_F$ (proof: see [4]).
- $\|A\|_2 = \sigma_{\max}(A) \leq \|A\|_{\mathrm{F}} \leq \sqrt{rank(A)}\|A\|_2$
- Consider a matrix norm $||.||$ and an $m \times m$ matrix A such that $||I_m|| = 1, ||A|| < 1$. Then $I_m - A$ is invertible and

$$||(I_m - A)^{-1}|| \leq \frac{1}{1 - ||A||}$$

- $||A^{-1} - (A+B)^{-1}|| \leq ||A^{-1}||^2||B||\,||(I_m + A^{-1}B)^{-1}|| \leq \frac{||A^{-1}||^2||B||}{1-||A^{-1}B||}$
- For any square matrix $A, ||A|| = \sqrt{\lambda}$ where λ is the largest eigenvalue of A^*A (proof: see [1], section 6.9).
- Let A be an invertible matrix. Then $||A^{-1}|| = \frac{1}{\sqrt{\lambda}}$, where λ is the smallest eigenvalue of A^*A (proof: see [1], section 6.9)

7 Block matrices

If a matrix is partitioned as

$$S = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \tag{1}$$

where A, B, C and D have arbitrary size, A and D are square matrices. If A and $D - CA^{-1}B$ must be nonsingular then,

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ -(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{bmatrix},$$

Equivalently, if $\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}$ is invertible,

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & -(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix}.$$

If we partition a matrix S as in 1 then

- $|S| = |D||A - BD^{-1}C|$ if D is nonsingular,
- $|S| = |A||D - CA^{-1}B|$ if A is nonsingular,
- $|S| \leq |A||D|$ if S is positive definite
- $rank(S) = rank(D) + rank(A - BD^{-1}C)$ if D is nonsingular,
- $rank(S) = rank(A) + rank(D - CA^{-1}B)$ if A is nonsingular,

8 Rank and matrix inversion

Rank properties from [3], section 2.4,

1. Let A be an $m \times n$ matrix. Then,
   - if B is an $n \times p$ matrix then,
   $$rank(A) + rank(B) - n \leq rank(AB) \leq \min\{rank(A), rank(B)\},$$
   - $|rank(A) - rank(B)| \leq rank(A + B) \leq rank(A) + rank(B)$ if B is an $m \times n$ matrix,
   - $rank(A) = rank(A') = rank(AA') = rank(A'A)$.
2. Let A, B, C be $p \times m, m \times n, n \times q$ matrices, respectively. Then,
   $$rank(ABC) \geq rank(AB) + rank(BC) - rank(B)$$

   We also have

1. $rank(A) = rank(AB)$ if A is an $m \times n$ matrix and B is an $n \times p$ matrix with $rank(B) = n$.

The Woodbury matrix identity:

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U \left(C^{-1} + VA^{-1}U\right)^{-1} VA^{-1},$$

where A, U, C and V all denote matrices of conformable sizes. This leads to the following

1. $(I + UV)^{-1} = I - U(I + VU)^{-1} V$.
2. $(I + P)^{-1} = I - (I + P)^{-1}P = I - P(I + P)^{-1}$
3. Push-through identity: $(I + UV)^{-1}U = U(I + VU)^{-1}$.
4. $$\begin{aligned}(A + B)^{-1} &= A^{-1} - A^{-1}(B^{-1} + A^{-1})^{-1}A^{-1} \\ &= A^{-1} - A^{-1}\left(I + BA^{-1}\right)^{-1} BA^{-1}.\end{aligned}$$
5. Hua identity: $(A + B)^{-1} = A^{-1} - \left(A + AB^{-1}A\right)^{-1} = A^{-1} + A^{-1}B(A - B)^{-1}$
6. *Sherman - Morrison formula:* For an invertible square matrix $A \in \mathbb{R}^{n\times n}$ and column vectors $u, v \in \mathbb{R}^n$, $A + uv^T$ is invertible iff $1 + v^TA^{-1}u \neq 0$. In this case,
$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u}.$$
7. If A, U, B, V are matrices of sizes $p \times p, p \times q, q \times q, q \times p$, respectively, then
$$(A + UBV)^{-1} = A^{-1} - A^{-1}UB\left(B + BVA^{-1}UB\right)^{-1} BVA^{-1}$$
provided A and $B + BVA^1UB$ are nonsingular. Nonsingularity of the latter requires that B^1 exist since it equals $B(I + VA^1UB)$ and the rank of the latter cannot exceed the rank of B.
8. Binomial Inverse Theorem: For B not neccesarily square or invertible,
$$(A + UBV)^{-1} = A^{-1} - A^{-1}\mathbf{U}(I + BVA^{-1}U)^{-1}BVA^{-1}.$$

Searle set of identities: (see [3])

1. $(I + A^{-1})^{-1} = A(A + I)^{-1}$
2. $(A + BB^T)^{-1}B = A^{-1}B(I + B^TA^{-1}B)^{-1}$
3. $(A^{-1} + B^{-1})^{-1} = A(A + B)^{-1}B = B(A + B)^{-1}A$
4. $A - A(A + B)^{-1}A = B - B(A + B)^{-1}B$
5. $A^{-1} + B^{-1} = A^{-1}(A + B)B^{-1}$
6. $(I + AB)^{-1} = I - A(I + BA)^{-1}B$

7. $(I + AB)^{-1}A = A(I + BA)^{-1}$

Other properties:

1. If $(A + B)^{-1} = A^{-1} + B^{-1}$ then $AB^{-1}A = BA^{-1}B$ (see [3])
2. When A has all eigenvalues λ_i satisfy $|\lambda_i| < 1$ then
   - $(I - A)^{-1} \approx I + A + A^2$
   - $(I + A)^{-1} \approx I - A + A^2$

9 Matrix decomposition

- **Cholesky decomposition**: Every real nonnegative-definite symmetric matrix A can be decompose as
$$A = LL^{\mathsf{T}}$$
where L is an $m \times m$ lower triangular matrix whose diagonal elements are nonnegative. If A is positive definite then L is unique and has positive diagonal elements.

- Let A be an $m \times n$ nonnegative definite matrix with $rank(A) = r$ then there exists an $m \times r$ matrix B of rank r such that $A = BB^T$

- Suppose that B is a $m \times h$ matrix and C is an $m \times n$ matrix, where $h \leq n$. Then, $BB^T = CC^T$ if and only if there exists an $h \times n$ matrix Q such that $QQ^T = I_h$ and $C = BQ$.

- **LU decomposition** Let A be an $m \times m$ matrix with non-zero leading principal minors, then $A = LU$ where L is a unique lower triangular matrix and U is a unique upper triangular matrix.

- **QR decomposition** Let A be an $m \times n$ matrix where $m \geq n$. Then there exists an $n \times n$ upper triangular matrix R with nonnegative diagonal elements and an $m \times n$ orthogonal matrix Q ($Q^TQ = I$) such that $A = QR$.

- **SVD decomposition** Suppose M is an $m \times n$ real or complex matrix. Then the singular value decomposition of M exists,
$$\mathbf{M} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*$$
where
  - U is an $m \times m$ unitary matrix,
  - Σ is a diagonal $m \times n$ matrix with non-negative real numbers on the diagonal. The diagonal entries σ_i of Σ are known as the singular values of M. If we list the singular values in descending order then Σ, is uniquely determined by M.
  - V is an $n \times n$ unitary matrix over K.

- Assume $V_* D_* U_*^T = 0$ then
$$A = [V \;\; V_*] \begin{pmatrix} D & 0 \\ 0 & D_* \end{pmatrix} \begin{pmatrix} U^T \\ U_*^T \end{pmatrix}$$
where $A = VDU^T$ is the SVD of A

- **Spectral decomposition and square root matrix:** Let A be a square $m \times m$ matrix with eigenvalue $\lambda_1, ..., \lambda_m$ and suppose that the corresponding set of orthornormal eigenvectors are $v_1, ..., v_m$. If $\Lambda = diag(\lambda_1, ..., \lambda_m)$, $Q = (v_1, ..., v_m)$ then,
$$\mathbf{A} = Q\Lambda Q^{-1} = \sum_{i=1}^{m} \lambda_i v_i v_i'.$$
Note that $QQ^T = Q^T Q = I$ and
$$A^{-1} = Q\Lambda^{-1} Q^T = \sum_{i=1}^{m} \frac{1}{\lambda_i} v_i v_i'.$$
Let $\Lambda^{1/2} = diag(\sqrt{\lambda_1}, ..., \sqrt{\lambda_m})$. Then the *square root* of A is defined as $A^{1/2} = Q\Lambda^{1/2} Q^T = \sum_{i=1}^{m} \sqrt{\lambda_i} v_i v_i'$ has the following properties
   1. $A^{1/2}$ is symmetric and $(A^{1/2})^2 = A$,
   2. $(A^{1/2})^{-1} = Q\Lambda^{-1/2} Q^T = \sum_{i=1}^{m} \frac{1}{\sqrt{\lambda_i}} v_i v_i'$. where $\Lambda^{-1/2} = diag(\frac{1}{\sqrt{\lambda_1}}, ..., \frac{1}{\sqrt{\lambda_m}})$,
   3. $A^{1/2} A^{-1/2} = A^{-1/2} A^{1/2} = I$ and $A^{-1/2} A^{-1/2} = A^{-1}$ where $A^{-1/2} = (A^{1/2})^{-1}$.

- **Schur decomposition:** Let A be a $m \times m$ matrix. Then there exists an $m \times m$ unitary matrix X such that
$$X^* A X = T$$
where T is an upper triangular matrix with the eigenvalues of A as its diagonal elements.

- Let A be a $m \times m$ matrix with real eigenvalues. Then there exists an $m \times m$ orthogonal matrix X such that
$$X^T A X = T$$
where T is an upper triangular matrix.

- The **polar decomposition** of a square complex matrix A is a matrix decomposition of the form
$$A = UP$$
where U is a unitary matrix and P is a positive-semidefinite Hermitian matrix.

- If A is $m \times m$ positive semi-definite, then there exists an $m \times r$ matrix B of rank r such that $B^T A B = I$

- $A = \frac{1}{2}(A + A^T) - \frac{1}{2}(A^T - A)$ and we have $A + A^T$ is symmetric and $A^T - A$ is anti-symmetric.

10 Special kind of matrices

10.1 Symmetric matrix

1. The sum and difference of two symmetric matrices is again symmetric.
2. Given symmetric matrices A and B, then AB is symmetric if and only if $AB = BA$.
3. For $n \in \mathbb{R}$, A^n is symmetric if A is symmetric.
4. If A^{-1} exists, it is symmetric if and only if A is symmetric.
5. Let $A \in \mathbb{R}^{m\times m}$ be a symmetric matrix. Then the eigenvalues of A are real, and corresponding to any eigenvalue, eigenvectors that are real exist.
6. Let $A \in \mathbb{R}^{m\times m}$ be a symmetric matrix. Suppose that A has r nonzero eigenvalues. Then, $rank(A) = r$.
7. Let $A \in \mathbb{R}^{m\times m}$ be a symmetric matrix with eigenvalue $\lambda_1, ..., \lambda_n$.
8. If A is symmetric then $x'Ax = 0 \;\; \forall x$ if and only if $A = 0$

10.2 Positive/semipositive/negative definite matrix

In this section, we denote λ_j as the jth largest eigenvalue of a matrix.

1. If A is positive definite then its diagonal entries are positive.
2. If A is positive definite then $rank(BAB^T) = rank(B)$
3. If A is $m \times m$ positive definite, B is $r \times m$ of rank r then BAB^T is positive definite.
4. If A is $n \times r, n \leq r$ and $rank(A) = n$ then AA^T is positive definite.
5. If A and B are positive definite, then so is A+B. The matrix inverse of a positive definite matrix is also positive definite.

   Moreover, $A - B$ is positive definite if and only if $B^{-1} - A^{-1}$ is positive definite.
6. If A is an $m \times m$ symmetric matrix, and A_k is its leading $k \times k$ principal submatrix. Then, A is positive definite if and only if all its leading principal minors, $A_1, ..., A_m$ are positive definite.
7. Let A be an $m \times m$ positive definite matrix, then

   $$A_\alpha = A - \alpha e_1 e_1',$$

   where $\alpha = |A|/|A_1|$ and A_1 is the $(m-1) \times (m-1)$ submatrix of A formed by deleting its first row and column. Then A_α is nonnegative definite.

8. Let A, B be $m \times m$ symmetric matrices. If A, B are nonnegative definite, then the ith largest eigenvalue of $A \odot B$ satisfies

$$\lambda_m(A) \left\{ \min_{1 \le i \le m} b_i i \right\} \le \lambda_i(A \odot B) \le \lambda_1(A) \left\{ \max_{1 \le i \le m} b_i i \right\}$$

9. Let A be an $m \times m$ positive definite matrix. Then $(A \odot A^{-1}) - I_m$ nonnegative definite.

10. Let A, B be $m \times m$ nonnegative definite matrices. Then

$$\lambda_m(A \odot B) \ge \lambda_m(AB).$$

11. Let $\lambda_1, .., \lambda_m$ be eigenvalues of an $m \times m$ symmetric matrix A. Then A is positive semidefinite if $\lambda_i \ge 0 \; \forall 1 \le i \le m$, and is positive definite if $\lambda_i > 0 \; \forall 1 \le i \le m$.

12. Let A be an $m \times n$ matrix with rank(A) = r. Then $A'A$ has r positive eigenvalues. Then,
    - It is positive definite if $r = n$ and positive semidefinite if $r < n$.
    - The positive eigenvalues of $A'A$ are equal to the positive eigenvalues of AA'

13. Let A be an $m \times m$ symmetric matrix and B be an $m \times m$ nonnegative definite matrix. Then, for $h = 1, ..., m$,

$$\lambda_h(A + B) \ge \lambda_h(A),$$

where the inequality is strict if B is positive definite.

14. Let A be an $m \times m$ symmetric matrix and B be an $m \times m$ positive definite matrix. If F is any $m \times h$ matrix with full column rank, then for $i = 1, ..., h$,

$$\lambda_i((F'BF)^{-1}(F'AF)) \le \lambda_i(B^{-1}A),$$

and

$$\max_F \lambda_i((F'BF)^{-1}(F'AF)) = \lambda_i(B^{-1}A),$$

10.3 Idempotent matrices and Hat matrices

A matrix A is idempotent if $A^2 = A$. Then

- $I - A, A^T, I_A^T$ are all idempotent
- AB is idempotent if $AB = BA$
- $rank(A) = trace(A)$
- $A(I - A) = (I - A)A = 0$

11 Special matrix operation

11.1 The vec operator

Vec operator is the operator that transforms a matrix into a vector: Suppose that we have a matrix $A = [a_1|a_2|...|a_n]$ where a_i is the i^{th} column of A. Then

$$vec(A) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ v_n \end{bmatrix}$$

Properties:

1. Let a, b be any 2 vectors and A, B be two matrices of the same size then
   - $vec(a) = vec(a') = a$
   - $vec(ab') = vec(b_1 a, b_2 a, ..., b_n a]) = \begin{bmatrix} b_1 a \\ \vdots \\ b_n a \end{bmatrix} = b \otimes a$
   - $vec(\alpha A + \beta B) = \alpha\, vec(A) + \beta\, vec(B)$ where $\alpha, \beta \in \mathbb{R}$
   - $trace(A'B) = \{vec(A)\}' vec(B)$
2. Let A, B, C, D be matrices of sizes $m \times n, n \times p, p \times q, q \times m$, respectively. Then
   - $vec(ABC) = (C' \otimes A) vec(B)$
   - $trace(ABCD) = \{vec(A')\}'(D' \otimes B) vec(C)$
3. Let A, C be matrices of sizes $m \times n, n \times m$, respectively, and B, D are marices of order $n \times n$. Then,
   - $trace(ABC) = \{vec(A')\}'(I_m \otimes B) vec(C)$
   - $trace(AD'BDC) = \{vec(D)\}'(A'C' \otimes B) vec(D)$

11.2 The Hadamard product

$$A \odot B = \begin{bmatrix} a_{11}b_{11} & ... & a_{1n}b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}b_{m1} & ... & a_{mn}b_{mn} \end{bmatrix}$$

Properties:

1. Let $a = (a_1, ..., a_n)', b = (b_1, ..., b_n)', c = (c_1, ..., c_n)'$ be vectors in $\mathbb{R}^n$. Then,

$$a'(b \odot c) = (a_1, .., a_n) \begin{pmatrix} b_1 c_1 \\ \vdots \\ b_n c_n \end{pmatrix} = a_1 b_1 c_1 + ... + a_n b_n c_n$$
$$= (a \odot b)'c = (a \odot c)'b = b'(a \odot c) = c'(a \odot b) = (c \odot a)'b$$

2. Let A, B, C be $m \times n$ matrices, D be an $m \times m$ diagonal matrix, and E be an $n \times n$ diagonal matrix; a, c are $m \times 1$ vectors and b, d are $n \times 1$ vectors. Then,

   (a) $A \odot B = B \odot A$,

   (b) $(A \odot B) \odot C = A \odot (B \odot C)$,

   (c) $(A + B) \odot C = A \odot C + B \odot C$,

   (d) $(A \odot B)' = A' \odot B'$,

   (e) $A \odot (0) = (0)$,

   (f) $A \odot 1_m 1'_n = A$,

   (g) $A \odot I_m = D_A = diag(a_{11}, ..., a_{mm})$ if $m = n$,

   (h) $D(A \odot B) = (DA) \odot B = A \odot (DB)$

   (i) $(A \odot B)E = (AE) \odot B = A \odot (BE)$

   (j) $ab' \odot cd = (a \odot c)(b \odot d)'$, where a, c,

   (k) $rank(A \odot B) \le rank(A) rank(B)$

   (l) $1'_m (A \odot B) 1_n = trace(AB')$

   (m) $a'(A \odot B)b = trace(D_a A D_b B')$ where $D_a = diag(a_1, ..., a_m), D_b = diag(b_1, ..., b_n)$.

3. Let A, B be $m \times m$ symmetric matrices. Then

   - If A, B both are nonnegative definite then $(A \odot B)$ is nonnegative definite,
   - If A, B both are positive definite then $(A \odot B)$ is positive definite,
   - If B is positive definite and A is nonnegative definite with positive diagonal elements, then $A \odot B$ is positive definite.

12 Matrix derivatives

kth-order Taylor formula

$$f(x + u) = f(x) + \sum_{i=1}^{k} \frac{u^i f^{(i)}(x)}{i!} + r_k(u, x)$$

where $\lim_{u \to} \frac{r_k(u,x)}{u^k} = 0$

If f is a real-valued function of $x = (x_1, ..., x_n)'$ then if its derivative at x exists, it is given by

$$\frac{\partial}{\partial x'} = \left[\frac{\partial}{\partial x_1} f(x), ..., \frac{\partial}{\partial x_n} f(x)\right].$$

First order Taylor formula

$$f(x+u) = f(x) + \left(\frac{\partial}{\partial x'} f(x)\right) u + r_1(u,x),$$

where $\lim_{u \to 0} \frac{r_1(u,x)}{(u'u)^{1/2}} = 0$.

kth-order Taylor formula

$$f(x+u) = f(x) + \sum_{i=1}^{k} \frac{d^i f}{i!} + r_k(u,x)$$

where $\lim_{u \to 0} \frac{r_k(u,x)}{(u'u)^{k/2}} = 0$.

We have $d^2 f = u'Hu$ where H is the Hessian matrix

$$H = \begin{bmatrix} \frac{\partial^2}{\partial x_1^2} f(x) & \cdots & \frac{\partial^2}{\partial x_1 \partial x_n} f(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial x_n \partial x_1} f(x) & \cdots & \frac{\partial^2}{\partial x_n^2} f(x) \end{bmatrix}$$

Suppose $f : \mathbb{R}^n \to \mathbb{R}^m$ is a function each of whose first-order partial derivatives exist on $\mathbb{R}^n$. Then the Jacobian matrix of f is defined as

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial x_1} & \cdots & \frac{\partial \mathbf{f}}{\partial x_n} \end{bmatrix}' = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix}.$$

First order Taylor formula

$$f(x+u) = f(x) + \left(\frac{\partial}{\partial x'} f(x)\right) u + r_1(u,x),$$

where $\lim_{u \to 0} \frac{r_1(u,x)}{(u'u)^{1/2}} = 0$.

Note that *if we obtain the first differential of f at x in u and $df = Bu$ then the $m \times n$ matrix B must be the derivative of f at x.*

Chain rule: If y and g are real-valued functions s.t. $y(x) = g(f(x))$ then

$$\frac{\partial}{\partial x'}y(x) = \left(\frac{\partial}{\partial f'}g(f)\right)\left(\frac{\partial}{\partial x'}f(x)\right),$$

$$\frac{\partial}{\partial x_i}y(x) = \sum_{j=1}^{m}\left(\frac{\partial}{\partial f_j}g(f)\right)\left(\frac{\partial}{\partial x_i'}f_j(x)\right)$$

$$= \left(\frac{\partial}{\partial f'}g(f)\right)\left(\frac{\partial}{\partial x_i'}f(x)\right)$$

For useful matrix derivative identities, see the matrix cookbook [5].

13 Convexity

A set S in $\mathbb{R}^m$ is **convex** if for any $x, y \in S$,

$$cx + (1-c)y \in S \qquad \forall 0 < c < 1$$

The intersection and union of two convex set are convex. The closure of a convex set is also convex.

From [3], section 2.11: Let S_1, S_2 be two convex set in $\mathbb{R}^m$ with $S_1 \cap S_2 = 0$. Then there exists $b \neq 0, b \in \mathbb{R}^{n\times 1}$ such that $b'x_1 \geq b'x_2 \; \forall x_1 \in S_1, x_2 \in S_2$.

References

[1] Stephen H. Friedberg, Arnold J. Insel, Lawrence E. Spence Linear Algebra (3rd Edition).Prentice Hall, 1996.

[2] Johnson, Richard Arnold, and Dean W. Wichern. Applied multivariate statistical analysis. Vol. 5. No. 8. Upper Saddle River, NJ: Prentice hall, 2002.

[3] Searle, Shayle R., and Andre I. Khuri. Matrix algebra useful for statistics. John Wiley Sons, 2017.

[4] Henry (https://math.stackexchange.com/users/193914/henry), Frobenius norm of product of matrix, URL (version: 2019-08-27): https://math.stackexchange.com/q/1393301

[5] Petersen, Kaare Brandt, and Michael Syskind Pedersen. "The matrix cookbook." Technical University of Denmark 7.15 (2008): 510.

[6] Wikipedia contributors. "Eigenvalues and eigenvectors." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 19 Sep. 2019. Web. 23 Sep. 2019.

[7] Wikipedia contributors. "Matrix norm." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 2 Sep. 2019. Web. 23 Sep. 2019.

[8] Wikipedia contributors. ”Block matrix.” Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 19 Sep. 2019. Web. 23 Sep. 2019.

[9] Wikipedia contributors. ”Singular value decomposition.” Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 18 Sep. 2019. Web. 23 Sep. 2019.

[10] Wikipedia contributors. ”ShermanMorrison formula.” Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 3 Jun. 2019. Web. 23 Sep. 2019.

[11] Wikipedia contributors. ”Symmetric matrix.” Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 18 Aug. 2019. Web. 23 Sep. 2019.

[12] Wikipedia contributors. ”Woodbury matrix identity.” Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 18 Sep. 2019. Web. 23 Sep. 2019.

[13] Schott, James R. Matrix analysis for statistics. John Wiley Sons, 2016.