M. Perlman

Some Useful Results from Matrix Theory

A. Notation and Conventions.

All vectors and matrices have real components. Vectors are column vectors. For any $p \times p$ square matrix $A = (a_{ij}), i,j = 1,\ldots,p$, the determinant of A is denoted by $|A|$ or $\det A$, and the trace of A by $\operatorname{tr}A$. I_p (or I) denotes the $p \times p$ identity matrix. $\operatorname{Diag}(a_1,\ldots,a_p)$ denotes the diagonal matrix with diagonal elements $a_1,\ldots,a_p$.

For any $p \times q$ rectangular matrix $B = (b_{ij})$, the transpose of B, denoted by B', is the $q \times p$ matrix whose ij-th element is b_{ji}. A square matrix B is symmetric if $B = B'$.

If $A(p \times p)$ is partitioned into blocks as

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1r} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rr} \end{pmatrix}$$

then $A_{ii\cdot j} = A_{ii} - A_{ij}A_{jj}^{-1}A_{ji}$ if A_{jj}^{-1} exists. (Whenever we consider a partitioned square matrix we assume that each diagonal block is a square matrix.)

The $(p \times p)$ matrix E_{ij} has a 1 in the ij-th position and zeroes everywhere else. Let $A_{ij} = I + E_{ij}$, and $M_i(c) = I + (c-1)E_{ii}$. A_{ij} and $M_i(c)$ are called elementary matrices.

A square matrix $T = (t_{ij})$ is lower (upper) triangular if $t_{ij} = 0$ for all $i < j$ $(i > j)$. If T is block partitioned as $T = (T_{ij})$, T is lower (upper) block triangular if $T_{ij} = 0$ for all $i < j$ $(i > j)$.

A principal submatrix of $A(p \times p)$ is a matrix obtained by deleting r rows and the corresponding r columns of A (any $r < p$). A leading principal submatrix is obtained by deleting the final r rows and columns. A (leading) principal minor is the determinant of a (leading) principal submatrix.

B. Positive (Semi-) Definite Matrices.

A $p \times p$ symmetric matrix Σ is called positive semi-definite (positive definite) if the quadratic form $x'\Sigma x \geq 0$ (> 0) for all non-zero vectors $x(p \times 1)$. $\Sigma \geq 0$ $(\Sigma > 0)$ denotes that Σ is positive semi-definite (positive definite). $A \geq B$ $(A > B)$ means that $A - B \geq 0$ (> 0).

1. For any $p \times q$ matrix A and any $q \times q$ $C \geq 0$, $ACA' \geq 0$. If $p \leq q$ and $C > 0$, $ACA' > 0$ iff A has full rank p.

2. $\Sigma > 0$ implies any principal submatrix of Σ is positive definite. $\Sigma > 0$ implies Σ^{-1} exists and $\Sigma^{-1} > 0$.

3. $\Sigma \geq 0$ (> 0) implies $|\Sigma| \geq$ (> 0). Converse is false.
$\Sigma \geq 0$ (> 0) iff each leading principal minor of Σ is ≥ 0 (> 0).

4. $A \geq B > 0$ implies (a) $|A| \geq |B|$
(b) $\mathrm{tr}A \geq \mathrm{tr}B$
(c) $B^{-1} \geq A^{-1}$
(Follows from E.3 below)
(d) $ch_i(A) \geq ch_i(B)$, $i = 1, \ldots, p$ (see E.2. below).
(Apply Courant-Fischer minmax Theorem)

C. <u>Partitioned Matrices</u>.

1. If $A(p \times p) = (A_{ij})$, $i, j = 1, 2$, then

$$|A| = |A_{11}|\ |A_{22} - A_{21}A_{11}^{-1}A_{12}| \text{ (if } A_{11}^{-1} \text{ exists).}$$
$$= |A_{22}|\ |A_{11} - A_{12}A_{22}^{-1}A_{21}| \text{ (if } A_{22}^{-1} \text{ exists).}$$

2. If $A > 0$ and $B = A^{-1}$, partition $A = (A_{ij})$ and $B = (B_{ij})$, $i, j = 1, 2$. Then the relation $AB = I$ implies

$$B_{11}^{-1} = A_{11\cdot 2}$$
$$B_{22}^{-1} = A_{22\cdot 1}$$
$$B_{12}A_{22\cdot 1} = -A_{11}^{-1}A_{12}$$
$$B_{21}A_{11\cdot 2} = -A_{22}^{-1}A_{21}\ .$$

Since $A > 0$ implies $B > 0$ implies $B_{ii} > 0$ implies $B_{ii}^{-1} > 0$, the first two relations imply $A_{ii\cdot j} > 0$, so the last two relations give

$$B_{12} = -A_{11}^{-1}A_{12}A_{22\cdot 1}^{-1}$$
$$B_{21} = -A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1} .$$

Since $B_{12} = B_{21}'$, we get

$$A_{12}A_{22\cdot 1}^{-1}A_{22} = A_{11}A_{11\cdot 2}^{-1}A_{12} .$$

D. Miscellaneous.

1. If A is $p \times q$ and B is $q \times p$, $\mathrm{tr}AB = \mathrm{tr}BA$. If $p = q$, $|AB| = |A|\cdot|B|$. $|A^{-1}| = |A|^{-1}$.

2. If $T(p \times p)$ is lower (upper) triangular then T^{-1} (if it exists) is also lower (upper) triangular. A similar statement holds for block triangular matrices. If A and B are lower (upper) triangular then AB is also lower (upper) triangular (similarly for block triangular matrices).

E. Matrix Representations and Factorizations.

1. $M_i(c)$ acting on the left (right) of a matrix A simply multiplies the i-th row (column) of A by the scalar c. A_{ij} acting on the left (right) of A simply adds the j-th row (column) to the i-th row (column). Then any nonsingular square matrix A can be expressed as a finite product of elementary matrices, i.e., A_{ij}'s and $M_i(c)$'s. That is, every such A can be reduced to the identity by

a finite number of elementary row and column operations.

2. If $\Sigma^{(p \times p)} \geq 0$ then Σ can be expressed as $\Gamma D_\lambda \Gamma'$, where $\Gamma(p \times p)$ is orthogonal and $D_\lambda = \text{diag}(\lambda_1, \ldots, \lambda_p)$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ are the characteristic roots of Σ, i.e., the roots of the equation $|\Sigma - \lambda I| = 0$. We use the notation $ch_i(\Sigma) = \lambda_i$. $\Sigma > 0$ iff $\lambda_p > 0$. If $\lambda_1 > \lambda_2 > \cdots > \lambda_p > 0$, then Γ is unique up to multiplication on the right by matrices of the form $\text{diag}(\pm 1, \ldots, \pm 1)$.

If $\Sigma \geq 0$ the symmetric square root of Σ, denoted by $\Sigma^{\frac{1}{2}}$, is defined as

$$\Sigma^{\frac{1}{2}} = \Gamma \, \text{diag}(\lambda_1^{\frac{1}{2}}, \ldots, \lambda_p^{\frac{1}{2}}) \, \Gamma' .$$

(Other powers of Σ are similarly defined.) Then $\Sigma^{\frac{1}{2}} \Sigma^{\frac{1}{2}} = \Sigma$.

If $\Sigma > 0$, let $\Sigma^{-\frac{1}{2}} = (\Sigma^{\frac{1}{2}})^{-1}$. Then $\Sigma^{-\frac{1}{2}} \Sigma \Sigma^{-\frac{1}{2}} = I$. Some authors define a square root of Σ to be any A such that $AA' = \Sigma$. (Note that A need not be symmetric, in fact not even a square matrix.) Any two square roots of Σ are related by a row-orthogonal transformation, as seen in 6 below.

3. Let $A^{p \times p} > 0$, $B^{p \times p} \geq 0$. Then there is a $p \times p$ nonsingular W such that $A = WW'$, $B = WD_\theta W'$, where

$D_\theta = \text{diag}(\theta_1,\ldots,\theta_p)$ where $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_p \geq 0$ are the roots of the equation $|B - \theta A| = 0$, or equivalently, the roots of $|A^{-1}B - \theta I| = 0$ or $|A^{-\frac{1}{2}}BA^{-\frac{1}{2}} - \theta I| = 0$. If in addition $B > 0$, then all $\theta_i > 0$. If in addition $\theta_1 > \theta_2 > \cdots \theta_p > 0$ then W is unique up to right multiplication by matrices of the form $\text{diag}(\pm 1,\ldots,\pm 1)$.

4. Let $A \geq 0$, $B \geq 0$. Then there exists an orthogonal matrix Γ such that $\Gamma A \Gamma'$ and $\Gamma B \Gamma'$ are diagonal, if and only if $AB = BA$.

5. If $\Sigma \geq 0$, there exists a lower (upper) triangular matrix T such that $\Sigma = TT'$ and $t_{ii} \geq 0$, $i=1,\ldots,p$. If $\Sigma > 0$ then then $t_{ii} > 0$ and T is unique.

6. (Vinograd's Theorem.) If A is $p \times r$ and B is $p \times s$ with $r \geq s$, then $AA' = BB'$ if and only if there exists a row-orthogonal matrix $\Gamma(s \times r)$, i.e. $\Gamma\Gamma' = I$, such that $A = B\Gamma$.

7. ("Woodbury's Theorem") If $A : p \times p$ and $S : q \times q$ are symmetric and U and V are $p \times q$, then

$$(A + USV')^{-1} = A^{-1} - A^{-1}U(S^{-1} + V'A^{-1}U)^{-1}V'A^{-1}$$

provided all inverses exist.

8. (Singular Value Decomposition). Let $A : p \times q$ be an arbitrary matrix and let $t = \min(p, q)$. The singular values $\sigma_1(A) \geq \cdots \geq \sigma_t(A) \geq 0$ of A are defined by

$$\sigma_i \equiv \sigma_i(A) = \left[ch_i(AA')\right]^{\frac{1}{2}} \equiv \left[ch_i(A'A)\right]^{\frac{1}{2}}, \quad 1 \leq i \leq t,$$

i.e., the square roots of the (non-trivial) characteristic roots of AA'. Then there exist column-orthogonal matrices $\Gamma: p \times t$ and $\Psi: q \times t$ (i.e., $\Gamma'\Gamma = I_t = \Psi'\Psi$) such that

$$A = \Gamma D_\sigma \Psi'.$$

9. (Gram-Schmidt) Let $\Gamma: p \times q$ be a column-orthogonal matrix ($p \geq q$). Then there exists $\Gamma_1 : p \times (p - q)$ so that the $p \times p$ matrix

$$(\Gamma \;\; \Gamma_1)$$

is orthogonal.

10. (Cauchy-Schwartz Inequalities):

(i) x, y $p \times 1$: $(x^t y)^2 \leq (x^t x)(y^t y)$ [Equality iff $x = cy$]

(ii) Σ $p \times p$, pd: $(x^t y)^2 \leq (x^t \Sigma x)(y^t \Sigma^{-1} y)$ [Equality iff $\Sigma x = cy$]

(iii) X, Y $p \times n$: $[\mathrm{tr}\, X^t Y]^2 \leq (\mathrm{tr}\, X^t X)(\mathrm{tr}\, Y^t Y)$ [Equality iff $X = cY$]

(iv) Σ $p \times p$, p.d.: $[\mathrm{tr}\, X^t Y]^2 \leq (\mathrm{tr}\, X^t \Sigma X)(\mathrm{tr}\, Y^t \Sigma^{-1} Y)$ [Equality iff $\Sigma X = cY$]

11. (Cauchy-Binet Theorem) $A : p \times r$, $B : r \times q$. Then for $1 \leq k \leq p \wedge q \wedge r$,

$$(AB)^{(k)} = A^{(k)} B^{(k)},$$

where $A^{(k)}$ denotes the <u>k-th compound matrix</u> of A, defined as follows:

$A^{(k)}$ is $\binom{p}{k} \times \binom{n}{k}$. The rows [columns] of $A^{(k)}$ are in 1-1 correspondence with the $\binom{p}{k}$ $\left[\binom{n}{k}\right]$ subsets of size k chosen from $1, \dots, p$ $[1, \dots, n]$. If α $[\beta]$ is such a subset, the α, β-entry of $A^{(k)}$ is the determinant of the α, β-submatrix of A.

(The Cauchy-Binet formula follows from the LaPlace expansion for a determinant.)

12. Minkowski Inequality for determinants:
Let A, B be positive definite $p \times p$ matrices. Then

$$|A+B|^{\frac{1}{p}} \geq |A|^{\frac{1}{p}} + |B|^{\frac{1}{p}}.$$

Hint: Represent $A = EE'$, $B = EDE'$, where $E^{p \times p}$ is nonsingular and $D = \mathrm{diag}(d_1, \dots, d_p)$, where $d_i = ch_i(A^{-1}B) > 0$. Also, use the binomial expansion & apply the Arithmetic Mean - Geometric Mean inequality.

13. Corollary 1. If $0 < \theta < 1$, $|\theta A + (1-\theta)B|^{\frac{1}{p}} \geq \theta |A|^{\frac{1}{p}} + (1-\theta)|B|^{\frac{1}{p}}$.
[Just apply 1. to θA and $(1-\theta)B$.]

14. Corollary 2. $|\theta A + (1-\theta)B| \geq |A|^{\theta} |B|^{1-\theta}$.

[Just apply Corollary 2 and the concavity of $\log(x)$.]

Corollary 2 states that $\log|A|$ is a concave function on the cone of $p \times p$ pos. def. matrices.