Lecture 0
Linear Algebra

Outline

1. Definitions of Matrices and Operations

- Definition: An $m \times n$ *matrix* A is a rectangular array of numbers with m rows and n columns. If the elements in the *i*th row and *j*th column is denoted a_{ij}, then A is often written $[a_{ij}]$, to be read "matrix whose *i, j*th element is a_{ij}.
- Definition: Two $m \times n$ matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are *equal*, written $A = B$, if $a_{ij} = b_{ij}$ for all *i* and *j.*
- Definition: The *Kronecker delta* δ_{ij} is defined by $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise.
- Definition: The $n \times n$ *identity matrix* I (or I_n) is defined by $I = [\delta_{ij}]$.
- Definition: The *null matrix* $\mathbf{0}$ is the matrix of all zero elements.
- Definition: An $n \times n$ matrix is *triangular* iff $a_{ij} = 0$ for $i > j$; and it is *diagonal* iff $a_{ij} = 0$ for $i \neq j$.

- Definition: Let A be a square matrix partitioned into blocks A_{ij} such that the diagonal blocks A_{ii} of $A = [A_{ij}]$ are square. Then A is *block triangular* iff $A_{ij} = 0$ for $i > j$; and A is *block diagonal* iff $A_{ij} = 0$ for $i \neq j$.
- Definition: The *transpose* A' of $A = [a_{ij}]$ is $A' = [a_{ji}]$.
- Definition: A matrix A is *symmetric* iff $A = A'$.
- Definition: Given $n \times n$ matrices A and B, B is called the *inverse* of A iff $AB = I$ and $BA = I$. When A has an inverse, it is typically written A^{-1}.
- Definition: If A^{-1} exists, A is said to be *nonsingular*; if A^{-1} does not exist, A is said to be *singular*.

- Definition: An $n \times n$ matrices A is *idempotent* iff $A^2 = A$.
- Definition: An $n \times n$ real matrices A is *orthogonal* iff $A'A = I$.
- Definition: Given a scalar λ, an $m \times n$ matrix $A = [a_{ij}]$, an $m \times n$ matrix $B = [b_{ij}]$, and an $n \times p$ matrix $C = [c_{ij}]$, the *sum* $A + B$, *scalar product* λA, and *matrix product* AC are defined by

$$A + B = [a_{ij} + b_{ij}], \quad \lambda A = [\lambda a_{ij}], \quad AC = [\sum_{k=1}^{n} a_{ik} c_{kj}].$$

2. Basic Properties

- $(\alpha + \beta)A = \alpha A + \beta A.$
- $\alpha(A + B) = \alpha A + \alpha B.$
- $(\alpha\beta)A = \alpha(\beta A).$
- $\alpha(AB) = (\alpha A)B = A(\alpha B).$
- If A is $n \times 1$, C is $1 \times n$, and $B = \beta$, then $ABC = \beta AC$.

- $A + B = B + A$.
- $(A + B) + C = A + (B + C)$.
- $(AB)C = A(BC)$.
- $A(B + C) = AB + AC$.
- $(A + B)C = AC + BC$.

- $IA = AI = A$.
- $A + 0 = 0 + A = A$.
- $A - A = 0$.
- $0A = A0 = 0$.
- $AB \neq BA$ is possible.
- $(A + B)' = A' + B'$.
- $(AB)' = B'A'$.
- $(A')' = A$.

- A^{-1} is unique.
- If A^{-1} exists and either $AB = I$ or $BA = I$, then $B = A^{-1}$ (B square).
- If A is symmetric and A^{-1} exists, A^{-1} is symmetric.
- $(AB)^{-1} = B^{-1}A^{-1}$.
- $(A')^{-1} = (A^{-1})'$.

- Suppose matrices A and B are partitioned into blocks A_{ij} and B_{ij} such that the columns of A are partitioned in the same way as the rows of B. Then

$$
\begin{aligned}
& AB \\
= & \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix} \begin{pmatrix} B_{11} & \cdots & B_{1p} \\ \vdots & & \vdots \\ B_{n1} & \cdots & B_{np} \end{pmatrix} \\
= & \begin{pmatrix} A_{11}B_{11} + \ldots + A_{1n}B_{n1} & \cdots & A_{11}B_{1p} + \ldots + A_{1n}B_{np} \\ \vdots & & \vdots \\ A_{m1}B_{11} + \ldots + A_{mn}B_{n1} & \cdots & A_{m1}B_{1p} + \ldots + A_{mn}B_{np} \end{pmatrix}
\end{aligned}
$$

3. Vector Spaces, Linear Dependence, Basis

- Definition (vector): An $n \times 1$ matrix is called an n-component *vector*.
- Definition (unit vector): The $n \times 1$ *unit vector* $\mathbf{e}_i$ is the $n \times 1$ vector with all components zero except the *i*th, which is 1.
- Definition (null vector): A *null vector* (or *zero vector*) $\mathbf{0}$ is a vector of all zero components.
- Definition (distance, length, angle): Let $\mathbf{a}$ and $\mathbf{b}$ be $n \times 1$ real vectors. *Distance*, *length*, and *angle* are defined by
  (Length of $\mathbf{a}$) $= |\mathbf{a}| = (\mathbf{a}'\mathbf{a})^{1/2}$
  (Distance from $\mathbf{a}$ to $\mathbf{b}$) $= |\mathbf{a} - \mathbf{b}| = [(\mathbf{a} - \mathbf{b})'(\mathbf{a} - \mathbf{b})]^{1/2}$
  (Angle between $\mathbf{a}$ and $\mathbf{b}$) $= \cos^{-1}(\mathbf{a}'\mathbf{b} \,/\, |\mathbf{a}|\,|\mathbf{b}|)$.

- Definition (orthogonal): Two n-component real vectors $\mathbf{a}$ and $\mathbf{b}$ are *orthogonal* if and only if $\mathbf{a}'\mathbf{b} = 0$.
- Definition (vector space): A non-empty set V of n-component vectors is a *vector space* if and only if it is closed under addition and multiplication by a scalar (that is, if and only if $\mathbf{a} + \mathbf{b}$ and $\lambda\mathbf{b}$ are in V when $\mathbf{a}$ and $\mathbf{b}$ are in V.)
- Definition (subspace): A *subspace* S of a vector space V is a subset of V which is itself a vector space.
- Definition (E^n): E^n denotes the set of all n-component vectors.

- Definition (linear dependence): A set of *n*-component vectors $\mathbf{a}_1, \ldots, \mathbf{a}_m$ is *linearly dependent* if and only if there exist scalar $\lambda_1, \ldots, \lambda_m$ not all zero such that $\lambda_1 \mathbf{a}_1 + \ldots + \lambda_m \mathbf{a}_m = 0$ (that is, such that one of the $\mathbf{a}_i$ can be expressed as a linear combination of the rest). Otherwise, the $\mathbf{a}_i$ are *linearly independent*.
- Definition (span): A set of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_r$ is said to *span* a vector space V if and only if every vector in V can be written as a linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_r$.

- Definition (basis): A *basis* for a vector space V is a subset of vectors in V which is linearly independent and spans V.
- Definition (orthogonal basis): An *orthogonal basis* for a real vector space is a basis of mutually orthogonal vectors.
- Definition (orthonormal basis): An *orthonormal basis* for a real vector space is an orthogonal basis of unit length vectors.
- Definition (dimension): The maximum number of linearly independent vectors in a vector space (when such a number exists) is called the *dimension* of V, written $\dim(V)$.

Theorem (length theorems): Let $\mathbf{a}$ and $\mathbf{b}$ be *n*-component real vectors. Then:

1. $|\mathbf{a}| + |\mathbf{b}| \geq |\mathbf{a} + \mathbf{b}|$ (triangle inequality).
2. $|\mathbf{a} + \mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2$ if $\mathbf{a}'\mathbf{b} = \mathbf{0}$ (Pythagoras theorem).
3. $|\mathbf{a}|\,|\mathbf{b}| \geq |\mathbf{a}'\mathbf{b}|$ (Schwartz inequality).

Theorem (basis changing theorem): If $\mathbf{a}_1, \ldots, \mathbf{a}_r$ is a basis for a vector space V, and if $\mathbf{b} = \lambda_1 \mathbf{a}_1 + \ldots + \lambda_r \mathbf{a}_r$ is a vector in V such that $\lambda_j \neq 0$, then the set $\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}$, $\mathbf{b}$, $\mathbf{a}_{j+1}, \ldots, \mathbf{a}_r$ is also a basis for V.

Theorem (properties of bases): Let V be a vector space of *n*-component vectors other than the set $\{0\}$. Then:

1. $\dim(V)$ exists, and $\dim(V) \leq n$.
2. Any linearly independent set of $\dim(V)$ vectors in V is a basis for V.
3. Every basis for V has $\dim(V)$ vectors.
4. Any set of $m < \dim(V)$ linearly independent vectors in V may be extended to form a basis for V.
5. If $\mathbf{a}_1, \ldots, \mathbf{a}_{\dim(V)}$ is a basis for V and $\mathbf{b}$ is a vector in V, the coefficients $\lambda_1, \ldots, \lambda_{\dim(V)}$ of $\mathbf{b} = \sum_i \lambda_i \mathbf{a}_i$ are unique.

Theorem (properties of E^n):

1. E^n is a vector space.
2. Every set of *n*-component vectors spans a subspace of E^n.
3. The set of unit vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ form a basis for E^n.
4. $\dim(E^n) = n$.

4. Row Reduction

Definition (reduced echelon matrix): A *reduced echelon matrix* is a matrix such that:

1. All non-zero rows precede all zero rows.
2. The first non-zero element in a non-zero row is a 1; and it appears in a column to the right of the first non-zero element in all preceding rows.
3. The first non-zero element in each non-zero row is the only non-zero element in its column.

Definition (elementary row operations): The three *elementary row operations* on a matrix are:

1. Interchanging the *i*th and *j*th rows.
2. Multiplying the *i*th row by a non-zero scalar c.
3. Adding the *i*th to the *j*th row.

Definition (elementary matrices): In the context of matrices with n rows, the *elementary row operation matrices*, or *elementary matrices* are:

1. $E_I(i,j) =$ the matrix gotten from I_n by interchanging its ith and jth rows.
2. $E_M(i,c) =$ the matrix gotten from I_n by multiplying its ith row by a non-zero scalar c.
3. $E_A(i,j) =$ the matrix gotten from I_n by adding its ith to its jth row.

Theorem (some basic row reduction rules):

1. Any elementary row operation on a matrix A may be accomplished by pre-multiplying A by the corresponding elementary matrix.
2. $E_I(i,j)^{-1} = E_I(i,j)$.
3. $E_M(i,c)^{-1} = E_M(i,1/c)$.
4. $E_A(i,j)^{-1} = E_M(i,-1)E_A(i,j)E_M(i,-1)$.
5. Any matrix may be transformed into a reduced echelon matrix by a finite sequence of elementary row operations.
6. The only non-singular reduced echelon matrix is I.

Theorem (row reduction for non-singular matrix): Let A be a non-singular matrix and let $E_1, \ldots, E_p$ be elementary matrices such that $E_p \cdots E_1 A = R$, where R is a reduced echelon matrix. Then:

1. $R = I$
2. $A^{-1} = E_p \cdots E_1$. Hence A^{-1} may be computed by performing the elementary row operations represented by the E_i on I.
3. $A = E_p^{-1} \cdots E_1^{-1}$. Hence any non-singular matrix may be represented as a product of elementary matrices.

5. Rank

Definition (rank): The *rank* $r(A)$ of a matrix A is the maximum number of linearly independent columns.

Theorem (some basic results on rank):

1. Let V_A be the vector space spanned by the columns of A. Then $\dim(V_A) = r(A)$.
2. If P and Q are non-singular, then $r(PAQ) = r(PA) = r(AQ) = r(A)$.
3. Given a matrix A and a corresponding reduced echelon matrix R, $r(A) = r(R) =$ number of non-zero rows in R.
4. $r(A) = r(A')$.
5. $r(AB) \leq \min(r(A), r(B))$.
6. An $n \times n$ matrix A is non-singular iff $r(A) = n$.

6. Inversion

Theorem: If A and B are $n \times n$ matrices such that $AB = I$ or $BA = I$, then A and B are non-singular.

Theorem (inverse of a partitioned matrix): If A and $E = D - CA^{-1}B$ are non-singular, then

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

Theorem (inverse of a block triangular matrix): Let T be a block triangular matrix with diagonal blocks $T_{11}, \ldots, T_{kk}$. Then:

1. T is non-singular iff $T_{11}, \ldots, T_{kk}$ are all non-singular.
2. If T is non-singular, T^{-1} is of the following form, where the upper right (asterisk) blocks are zero in the block diagonal case.

$$T^{-1} = \begin{bmatrix} T_{11}^{-1} & & & * \\ & T_{22}^{-1} & & \\ & & \ddots & \\ 0 & & & T_{kk}^{-1} \end{bmatrix}.$$

Definition (generalized inverse): A *generalized inverse* of a matrix A, denoted A^-, satisfies the following requirements:

1. $AA^-A = A$.
2. $A^-AA^- = A^-$.
3. A^-A is symmetric.
4. AA^- is symmetric.

Theorem (properties of a generalized inverse):

1. A^- exists.
2. A^- is unique.
3. If A is square and non-singular, $A^- = A^{-1}$.

7. Determinant, Cofactor, Adjoint, Trace

Definition (determinant): The *determinant* $\det(A)$ of an $n \times n$ matrix $A = [a_{ij}]$ is

$$\det(A) = \sum \binom{+}{-} a_{1i_1} a_{2i_2} \cdots a_{ni_n},$$

where the sum is taken over all permutations $(i_1, i_2, \ldots, i_n)$ of the integers $(1, 2, \ldots, n)$. A term is assigned a plus (minus) sign if $(i_1, i_2, \ldots, i_n)$ has an even (odd) number of inversions [number of pairs of elements of $(i_1, i_2, \ldots, i_n)$, not necessarily adjacent, for which a larger integer precedes a smaller one].

Theorem (some basic results on determinant): Let A and B be $n \times n$ matrices, except in part 16. Then:

1. $\det[E_I(i,j)A] = -\det(A)$.
2. $\det[E_M(i,c)A] = c\det(A)$.
3. $\det[E_A(i,j)A] = \det(A)$.
4. If two rows of A are identical, $\det(A) = 0$.
5. $\det(I) = 1$.
6. $\det[E_I(i,j)] = -1, \det[E_M(i,c)] = c, \det[E_A(i,j)] = 1$.
7. $\det(A) = 0$ iff A is singular.

8. $\det(AB) = \det(A)\det(B) = \det(BA)$.
9. $\det(A') = \det(A)$.
10. $\det(A^{-1}) = 1/\det(A)$.
11. $\det(A) = a_{11}\cdots a_{nn}$ if A is diagonal or triangular.
12. $\det(A) = \det(A_{11})\cdots\det(A_{kk})$ if A is block diagonal or block triangular with diagonal blocks $A_{11}, \cdots, A_{kk}$.
13. If A is non-singular, $\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det\begin{pmatrix} D & C \\ B & A \end{pmatrix} = \det(A)\det(D - CA^{-1}B)$.

14. Let $A = (A_1, \cdots, A_n)$ and let the *j*th column be given by a sum $A_j = \sum_k C_k$ of columns C_k. Then

$$\det(A) = \sum_k \det(A_1, \cdots, A_{j-1}, C_k, A_{j+1}, \cdots, A_n).$$

15. $\det(cA) = c^n \det(A)$.
16. For any $m \times n$ matrix A, $r(A)$ equals the size of the largest square submatrix of A with a non-zero determinant.

Definition (minor and cofactor): Let A be an $n \times n$ matrix. For $n > 1$, let A_{ij} be the $(n-1) \times (n-1)$ submatrix obtained by deleting the *i*th row and *j*th column from A. Then the scalar $M_{ij} = \det(A_{ij})$ is called the (*i, j*)th *minor* of A and the scalar $C_{ij} = (-1)^{i+j} M_{ij}$ is called the (*i, j*)th *cofactor* of A. A cofactor is a signed minor. For $n = 1$, the minor and the cofactor of a_{11} is one when $a_{11} \neq 0$, and the minor and the cofactor of a_{11} is zero when $a_{11} = 0$.

Definition (adjoint): Given a matrix A with cofactor C_{ij}, the *adjoint* A^+ of A is the matrix $A^+ = [a^+_{ij}]$ where $a^+_{ij} = C_{ji}$.

Theorem (computing the inverse):

1. $AA^+ = \det(A)I$ or $\sum_j a_{ij}C_{kj} = \delta_{ik}\det(A)$ $(i, k = 1, \ldots, n)$.
2. $A^+A = \det(A)I$ or $\sum_i a_{ij}C_{ik} = \delta_{jk}\det(A)$ $(j, k = 1, \ldots, n)$.
3. If A is non-singular, $A^{-1} = A^+/\det(A)$.

Definition (trace): The *trace* $tr(A)$ of an $n \times n$ matrix $A = [a_{ij}]$ is defined by $tr(A) = a_{11} + \cdots + a_{nn}$.

Theorem (some basic results on trace):

1. $tr(A+B) = tr(A) + tr(B)$.
2. $tr(A') = tr(A)$.
3. $tr(cA) = c[tr(A)]$.
4. $tr(AB) = tr(BA) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ji}$ for A $m \times n$ and B $n \times m$.

8. Linear Equation Systems

Theorem: Fundamental results about the linear equation system $A\mathbf{x} = \mathbf{b}$, where A is $m \times n$, are as follows.

1. If $m = n$ and A is non-singular, the unique solution for $\mathbf{x}$ is $\mathbf{x} = A^{-1}\mathbf{b}$.
2. Cramer's Rule: If $m = n$ and $A = (A_1, \cdots, A_n)$ is non-singular, the unique solution for x_i is given by

$$x_i = \det(A_1, \cdots, A_{i-1}, \mathbf{b}, A_{i+1}, \cdots, A_n) / \det(A).$$

3. A solution exists iff $r(A|\mathbf{b}) = r(A)$. If a solution exists, it is unique iff $r(A) = n$.

4. If $(A|\mathbf{b})$ can be transformed into $(C|\mathbf{d})$ by elementary row operations, then the systems $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ have identical solutions. Thus, $A\mathbf{x} = \mathbf{b}$ may be solved as follows.

   0.1 Row reduce $(A|\mathbf{b})$ to reduced echelon form $(R|\mathbf{c})$ and consider solutions to $R\mathbf{x} = \mathbf{c}$.

   0.2 If $r(R|\mathbf{c}) > r(R)$, there is no solution. If $r(R|\mathbf{c}) = r(R)$, a solution exist; let $k = r(R)$ and proceed.

   0.3 Reorder the columns of R and the elements of $\mathbf{x}$ so that $R\mathbf{x} = \mathbf{c}$ partitions as

$$\begin{pmatrix} I & R_0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{c}_0 \\ 0 \end{pmatrix} \quad \text{or} \quad \mathbf{x}_1 = \mathbf{c}_0 - R_0\mathbf{x}_2,$$

   where I is $k \times k$, R_0 is $k \times (n-k)$, $\mathbf{x}_1$ and $\mathbf{c}_0$ are $k \times 1$, and $\mathbf{x}_2$ is $(n-k) \times 1$.

   0.4 Generate all possible solutions $(\mathbf{x}_1', \mathbf{x}_2')$ by arbitrarily specifying $\mathbf{x}_2$ and computing $\mathbf{x}_1$.

5. If $\mathbf{b} = \mathbf{0}$, the set of solutions for $\mathbf{x}$ is a vector space (called the *null space*) of dimension $n - r(A)$.
6. Let N be the vector space of solution to $A\mathbf{x} = \mathbf{0}$ (N is the null space), and let $\mathbf{x} = \boldsymbol{\alpha}$ be a solution to $A\mathbf{x} = \mathbf{b}$. Then the set of all solutions to $A\mathbf{x} = \mathbf{b}$ is $\{\mathbf{x} |\ \mathbf{x} = \boldsymbol{\alpha} + \boldsymbol{\beta}$ with $\boldsymbol{\beta}$ in $N\}$.

7. For any vector space S in E^n, there is a matrix A such that S can be represented as the solution set of $A\mathbf{x} = \mathbf{0}$ (that is, as the null space of A).
8. Let all arrays be real, let V be the vector space spanned by the columns of A' (called the *row space*), and let N be the vector space of solutions to $A\mathbf{x} = \mathbf{0}$ (the null space). Then:
   0.1 Any vector in V is orthogonal to any vector in N.
   0.2 A basis for V together with a basis for N forms a basis for E^n.

9. Eigenvalues and Eigenvectors

- Definition: λ is an *eigenvalue* of an $n \times n$ matrix A iff $\det(A - \lambda I) = 0$.
- Definition: $\det(A - \lambda I)$ is called the *characteristic polynomial* of A, and $\det(A - \lambda I) = 0$ is called the *characteristic equation* of A.
- Definition: Given a matrix A and an eigenvalue λ of A, a non-zero vector $\mathbf{x}$ is an *eigenvector* of A corresponding to λ iff $A\mathbf{x} = \lambda\mathbf{x}$.

Theorem: Let A be an $n \times n$ matrix. Then:

1. The characteristic polynomial of A is an *n*th degree polynomial which may be written $\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda)$, where the λ_i depend on the elements of A. Therefore, the characteristic equation has the n solutions $\lambda_1, \lambda_2, \ldots, \lambda_n$. That is, A has n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. When A is real, complex eigenvalues must come in conjugate pairs.
2. A and A' have the same characteristic polynomial and the same eigenvalues.
3. For any $n \times n$ non-singular matrix P, A and $P^{-1}AP$ have the same characteristic polynomial and the same eigenvalues.
4. $\det(A) = \lambda_1 \cdots \lambda_n$ and $tr(A) = \lambda_1 + \cdots + \lambda_n$, where λ_i are the eigenvalues of A.

5. A is non-singular iff all of its eigenvalues are non-zero.
6. If A is non-singular, the eigenvectors of A^{-1} are the same as those of A; and the eigenvalues of A^{-1} are the inverses of those of A.
7. Is A is real and symmetric, its eigenvalues are real; and its rank equals the number of non-zero eigenvalues.
8. If A is triangular, its diagonal elements are its eigenvalues. If A is block triangular, the eigenvalues of the diagonal blocks are the eigenvalues of A.

10. Diagonalization and Triangularization

Theorem: Some basic results are:

1. (**diagonalization**) Given any $n \times n$ real symmetric matrix A and its eigenvalues $\lambda_1, \ldots, \lambda_n$, there exists a real matrix Q such that: (a) $Q'AQ = diag(\lambda_1, \ldots, \lambda_n)$; (b) $Q'Q = I$;and (c) the columns of Q are eigenvectors of A corresponding to $\lambda_1, \ldots, \lambda_n$.
2. (**triangularization**) Given any $n \times n$ matrix A, there exists a non-singular matrix P such that $P^{-1}AP$ is triangular with the eigenvalues of A on its diagonal.

11. Quadratic Forms and Definiteness

Definition (quadratic form): A *quadratic form* is a scalar quantity of the form $\mathbf{x}'A\mathbf{x}$ where A is an $n \times n$ real symmetric matrix and $\mathbf{x}$ is an $n \times 1$ real vector.

Definition (definiteness): A real symmetric matrix A is:

1. *positive definite* iff $\mathbf{x}'A\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$,
2. *positive semidefinite* iff $\mathbf{x}'A\mathbf{x} \geq 0$ for all $\mathbf{x} \neq \mathbf{0}$,
3. *negative definite* iff $\mathbf{x}'A\mathbf{x} < 0$ for all $\mathbf{x} \neq \mathbf{0}$,
4. *negative semidefinite* iff $\mathbf{x}'A\mathbf{x} \leq 0$ for all $\mathbf{x} \neq \mathbf{0}$, and
5. *indefinite* otherwise.

Theorem: Let A be a real symmetric $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then:

1. A is

$$\begin{array}{rcl} \text{positive definite} & \text{iff} & \text{all } \lambda_i > 0, \\ \text{positive semidefinite} & \text{iff} & \text{all } \lambda_i \geq 0, \\ \text{negative definite} & \text{iff} & \text{all } \lambda_i < 0, \\ \text{negative semidefinite} & \text{iff} & \text{all } \lambda_i \leq 0, \text{ and} \\ \text{indefinite} & & \text{otherwise.} \end{array}$$

2. For any real non-singular $n \times n$ matrix R, $R'AR$ and A have the same type of definiteness.

3. If A is positive (negative) definite, then so is any submatrix of A gotten by deleting rows and corresponding columns.
4. Let B be an $n \times k$ real matrix, and let C be a $k \times k$ real symmetric matrix. If A is positive (negative) definite, then

$$\left\{ \begin{pmatrix} A & B \\ B' & C \end{pmatrix} \text{ is pos (neg) definite} \right\}$$
$$\iff \left\{ C - B'A^{-1}B \text{ is pos (neg) definite} \right\}.$$

5. For all real $n \times 1$ vectors $\mathbf{x}$ of unit length,

$$\max(\mathbf{x}'A\mathbf{x}) = \max(\lambda_1, \ldots, \lambda_n),$$

$$\min(\mathbf{x}'A\mathbf{x}) = \min(\lambda_1, \ldots, \lambda_n).$$

6. Let A_i be the submatrix of A gotten by deleting the last $n-i$ rows and columns of A. Then A is:

   positive definite iff $\det(A_i) > 0$ for all i, and

   negative definite iff $(-1)^i \det(A_i) > 0$ for all i.

7. Let B be a real symmetric $n \times n$ matrix. If $(A-B)$ is positive definite, then so is $(B^{-1} - A^{-1})$.

Theorem (cross products matrix $X'X$): Let X be an $n \times k$ real matrix. Then:

1. $X'X$ is symmetric.
2. $r(X'X) = r(X) = r(XX')$.
3. $X'X$ is positive definite iff $r(X) = k$, and positive semidefinite iff $r(X) < k$.

Definition (idempotent): An an $n \times n$ matrix A is *idempotent* if $A^2 = A$.

Theorem (properties of an idempotent matrix): Let A be an $n \times n$ real symmetric idempotent matrix. Then:

1. If A is non-singular, $A = I$.
2. $\det(A)$ is either 0 or 1.
3. Any eigenvalue of A is either 0 or 1.
4. $tr(A) = r(A)$.
5. If A is real, singular, and non-zero, then there exists a matrix Q such that $Q'Q = I$ and $Q'AQ = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$.

12. Kronecker Products and the Stacking Operator

Definition: Let $A^{m\times n} = [a_{ij}] = (A_1, \ldots, A_n)$ and B be any two matrices. Then the *Kronecker product* of A and B, written $A \otimes B$, and the *stack* of A,written $vec(A)$, are defined by

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \quad \text{and} \quad vec(A) = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix}.$$

Theorem:

1. $vec(A+B) = vec(A) + vec(B)$.
2. $vec(AYB) = (B' \otimes A)vec(Y)$.
3. $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.
4. $A \otimes (B \otimes C) = (A \otimes B) \otimes C$.
5. $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$.
6. $(B + C) \otimes A = (B \otimes A) + (C \otimes A)$.

7. $(A \otimes B)' = A' \otimes B'$.
8. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.
9. Let λ and μ be eigenvalues of square matrices A and B,respectively, with corresponding eigenvectors $\mathbf{x}$ and $\mathbf{y}$. Then $\lambda\mu$ is an eigenvalue of $A \otimes B$ with corresponding eigenvector $\mathbf{x} \otimes \mathbf{y}$.
10. $\det(A \otimes B) = [\det(A)]^n [\det(B)]^m$, where A is $m \times m$ and B is $n \times n$.
11. $tr(A \otimes B) = tr(A)tr(B)$.
12. $r(A \otimes B) = r(A)r(B)$.

13. Matrix Norms

Definition (matrix norm): A *matrix norm* is a real-valued scalar function f of a square matrix such that f satisfies the following four axioms, where A and B are any $n \times n$ matrices.

1. $f(A) \geq 0$ with equality holding iff $A = 0$.
2. $f(cA) = |c|\, f(A)$ for any scalar c.
3. $f(AB) \leq f(A)f(B)$.
4. $f(A + B) \leq f(A) + f(B)$.

Theorem (examples of norms): Let $A = [a_{ij}]$ be an $n \times n$ matrix. The following functions of A are all matrix norms.

1. Maximum element norm: $n \max_{i,j} |a_{ij}|$.
2. Holder norm: $[\sum_i \sum_j |a_{ij}|^q]^{1/q}$ for $1 \leq q \leq 2$.
3. Euclidean norm: $[\sum_i \sum_j |a_{ij}|^2]^{1/2}$ (Holder norm with $q = 2$).
4. Element sum norm: $\sum_i \sum_j |a_{ij}|$ (Holder norm with $q = 1$).
5. Column sum norm: $\max_j \sum_i |a_{ij}|$.
6. Row sum norm: $\max_i \sum_j |a_{ij}|$.
7. Weighted norm: $f(P^{-1}AP)$ where f is any matrix norm and P is any nonsingular matrix.

Theorem (norms and eigenvalues): Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$; and let f be any matrix norm function. Then $|\lambda_i| \leq f(A)$ for $i = 1, \ldots, n$.

14. Matrix Power and Matrix exponential

Theorem: The following statements about an $n \times n$ matrix A are equivalent:

1. For integer t, $A^t \to 0$ as $t \to \infty$.
2. The eigenvalues of A are all less than one in modulus.
3. The series $I + A + A^2 + \cdots$ converges and equals $(I - A)^{-1}$.
4. $T^t \to 0$ as $t \to \infty$, where $T = P^{-1}AP$ is a triangularized matrix.

Definition: The matrix exponential e^A is defined by the infinite series $e^A = \sum_{j=0}^{\infty} A^j/j!$

Theorem: Let A and B be $n \times n$ matrices. Then:

1. The series e^A converges for any A.
2. If $AB = BA$, then $e^{A+B} = e^A e^B$.
3. $(e^A)^{-1} = e^{-A}$.
4. λ is an eigenvalue of A if and only if e^λ is an eigenvalue of e^A.
5. $\det(e^A) = e^{tr(A)}$.
6. e^A is non-singular for any A.
7. For real t, $e^{At} \to 0$ as $t \to \infty$ iff all eigenvalues of A have negative real parts.

15. Differentiating Matrix Expressions

Definition:

1. $\partial A/\partial z = [\partial a_{ij}/\partial z]$.
2. $\partial z/\partial A = [\partial z/\partial a_{ij}]$.
3. $\partial \mathbf{y}/\partial \mathbf{x} = [\partial y_i/\partial x_j]$.

Theorem:

1. $\partial(A\mathbf{x})/\partial\mathbf{x} = A$.
2. $\partial(AB)/\partial z = (\partial A/\partial z)B + A(\partial B/\partial z)$.
3. $\partial(A \otimes B)/\partial z = (\partial A/\partial z) \otimes B + A \otimes (\partial B/\partial z)$.
4. $\partial(\mathbf{x}'A\mathbf{x})/\partial\mathbf{x} = (A + A')\mathbf{x} = 2A\mathbf{x}$ (with A symmetric).
5. $\partial(\mathbf{x}'A\mathbf{x})/\partial A = \mathbf{x}\mathbf{x}'$.
6. $\partial \ln |A|/\partial A = (A')^{-1}$.
7. $\partial A^{-1}/\partial z = -A^{-1}(\partial A/\partial z)A^{-1}$.
8. $\partial[\det(A)]/\partial z = tr[A^{+}(\partial A/\partial z)]$.
9. $(d/dt)e^{At} = Ae^{At}$.