Mathematics and Statistics Summary Sheets

Erik Erhardt

November 10, 2004

Contents

1 Linear Algebra Matrix Operations

1.1 Basic Operations

$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$.
$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$.
$\mathbf{AB} \neq \mathbf{BA}$.
$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

Elementary Row and Column Operations:

E1 Interchange any two rows.
E2 Multiply the elements of any row by a nonzero scalar.
E3 Add to any row, element by element, a scalar times the corresponding elements of another row.
rank: number of non-zero rows after row-echelon form.

1.2 Simultaneous Linear Equations

$\mathbf{A}\underset{\sim}{X} = \underset{\sim}{B}$.
Nilpotent: of index p, if $\mathbf{A}^p = 0$ for some $p \in \mathbb{Z}^+$.
Idempotent: if $\mathbf{A}^2 = \mathbf{A}$.
$(\mathbf{A}^p)^T = (\mathbf{A}^T)^p$.

1.3 Square Matricies

$\mathbf{I}$ is the diagonal identity matrix: $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$, $\mathbf{A}^0 = \mathbf{I}$.

1.4 Inverses

$[\mathbf{A}|\mathbf{I}] \to [\mathbf{I}|\mathbf{A}^{-1}]$.

1. Inverse of non-singular matrix is unique.
2. If non-singular, $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.
3. If $\mathbf{A}$, $\mathbf{B}$ are non-singular, $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
4. If $\mathbf{A}$ non-singular, then so is $\mathbf{A}^T$, and $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$.

1.5 Determinants

1. If $\mathbf{A}$, $\mathbf{B}$ same order, then $\text{Det}(\mathbf{AB}) = \text{Det}(\mathbf{A})\text{Det}(\mathbf{B})$.
2. Determinant of an upper or lower triangular matrix is the product of the diagonal elements.
3. Swap two rows or columns of $\mathbf{A}$ to get $\mathbf{B}$, then $\text{Det}(\mathbf{A}) = -\text{Det}(\mathbf{B})$.
4. Multiply a row (or column) of $\mathbf{A}$ by k to get $\mathbf{B}$, then $k\text{Det}(\mathbf{A}) = \text{Det}(\mathbf{B})$.
5. To $\mathbf{A}$, add a constant times one row (or column) to another row (or column) to get $\mathbf{B}$, then no change $\text{Det}(\mathbf{A}) = \text{Det}(\mathbf{B})$.
6. If $\mathbf{A}$ has a zero row (or column), then $\text{Det}(\mathbf{A}) = 0$.
7. If two rows are equal, $\text{Det}(\mathbf{A}) = 0$.
8. $\text{Det}(\mathbf{A}^T) = \text{Det}(\mathbf{A})$.
9. A non-square matrix $\mathbf{A}$ has **rank** k iff it possesses at least one $k \times k$ submatrix with non-zero Det while all larger submatricies have Det $= 0$.
10. If $\mathbf{A}$ has an inverse, $\text{Det}(\mathbf{A}^{-1}) = (\text{Det}(\mathbf{A}))^{-1}$

If $\text{Det}(\mathbf{A}) = 0$, $\mathbf{A}$ does not have an inverse.
$\text{Det}(k\mathbf{A}) = k^n\text{Det}(\mathbf{A})$.
$\text{Det}(\mathbf{A}^2) = (\text{Det}(\mathbf{A}))^2$.
$\text{Det}(\mathbf{AB}) = \text{Det}(\mathbf{BA})$.

1.6 Vectors

Linearly dependent: n m-dimensional vectors $\{\underset{\sim}{V}_1, \ldots, \underset{\sim}{V}_n\}$ are linearly dependent if there exist constants $c_1, \ldots, c_n$ not all zero such that $c_1V_1 + \cdots + c_nV_n = 0$.
Linearly independent: If $c_1, \ldots, c_n$ must all be zero for $c_1V_1 + \cdots + c_nV_n = 0$.
If the **rank** is smaller than the number of rows, the rows are linearly dependent, otherwise linearly independent.

Linearly dependent vector properties:

1. $m+1$ or more m-dimensional vectors are linearly dependent.
2. If a set of vectors is linearly independent then any subset are also linearly independent.
3. If a set is linearly dependent then any larger set is also linearly dependent.

The **rank** is also the maximum number of linearly independent vectors that can be formed from the row (or column) vectors. This is also called the row rank or column rank, but they are all the same.
The **null space** of matrix $\mathbf{A}$ are all $\underset{\sim}{X}$ with $\mathbf{A}\underset{\sim}{X} = 0$.

1.7 Eigen (Characteristic) values/vectors

$\underset{\sim}{X}$ is an eigenvector and λ is an eigenvalue when $\mathbf{A}\underset{\sim}{X} = \lambda\underset{\sim}{X}$.
λ may be 0, $\underset{\sim}{X}$ may not be $\underset{\sim}{0}$.
$\text{Det}(\mathbf{A} - \underset{\sim}{\lambda}\mathbf{I}) = 0$ is the characteristic polynomial.

Eigenvalue/vector properties:

1. Sum of eigenvalues equals the **trace**, the sum of the main diagonal.
2. Eigenvectors corresponding to different eigenvalues are linear independent.
3. Matrix $\mathbf{A}$ is **singular** iff $\mathbf{A}$ has a zero eigenvalue.
4. If $\underset{\sim}{X}$ is an eigenvector of $\mathbf{A}$ with eigenvalue λ and $\mathbf{A}$ is invertible, then $\underset{\sim}{X}$ is eigenvector of $\mathbf{A}^{-1}$ with eigenvalue λ^{-1}. (Also works for squares, but not sure if true in general.)
5. If $\underset{\sim}{X}$ is an eigenvector, then so is $k\underset{\sim}{X}$ for the same eigenvalue λ.
6. A matrix and its transpose, $\mathbf{A}$ and $\mathbf{A}^T$, have the same eigenvalues, $\underset{\sim}{\lambda}$.
7. The eigenvalues of an upper or lower triangular matrix are the elements on its main diagonal.
8. The product of eigenvalues are the determinant.
9. If $\underset{\sim}{X}$ is the eigenvector of $\mathbf{A}$ for eigenvalue λ, then $\underset{\sim}{X}$ is the eigenvector of $\mathbf{A} - c\mathbf{I}$ for eigenvalue $\lambda - c$ for any scaler c.

Determining eigenvalues/vectors:

1. $n \times n$ matrix $\mathbf{A} - \underset{\sim}{\lambda}\mathbf{I}$.
2. Simplify to one matrix.
3. Take determinant to get n-degree polynomial in terms of λ, characteristic equation.
4. Solve for λ for eigenvalues.
5. Get associated eigenvectors for $\underset{\sim}{\lambda}$ by:
6. $(\mathbf{A} - \lambda_1\mathbf{I})\underset{\sim}{X}_1 = \underset{\sim}{0}$ and solve for $\underset{\sim}{X}_1$.

7. This will be in terms of one of the elements of $\underset{\sim}{X}_1$ (some multiple).

1.8 Functions of Matrices

Matrix Equation

$\mathbf{A}\underset{\sim}{X}+\underset{\sim}{X}\mathbf{B}=\mathbf{C}$ has a unique solution iff square matricies $\mathbf{A}$,$\mathbf{B}$ have no eigenvalues in common.
Unique solution is $\underset{\sim}{X}=-\int_0^\infty e^{\mathbf{A}t}\mathbf{C}e^{\mathbf{B}t}\,\mathrm{d}t$.
$e^{\mathbf{A}t}e^{\mathbf{B}t}=e^{(\mathbf{A}+\mathbf{B})t}$ iff $\mathbf{AB}=\mathbf{BA}$, they commute.

1.9 Canonical Bases

Generalized Eigenvectors

Vector $\underset{\sim}{X}_m$ is a **Generalized (right) eigenvecotr of rank** m for square matrix $\mathbf{A}$ associated with eigenvalue λ if $(\mathbf{A}-\lambda\mathbf{I})^m\underset{\sim}{X}_m=0$ but $(\mathbf{A}-\lambda\mathbf{I})^{m-1}\underset{\sim}{X}_m\neq 0$.
These are usually of rank 1.

Chains

A **chain** generated by a generalized eigenvector $\underset{\sim}{X}_m$ of rank m associated with eigenvalue λ is a set of vectors $\{\underset{\sim}{X}_m, X_{m-1}, \ldots, \underset{\sim}{X}_1\}$ defined recursively as $\underset{\sim}{X}_j=(\mathbf{A}-\lambda\mathbf{I})\underset{\sim}{X}_{j+1}$, $(j=m-1,\ldots,1)$.
A chain is a linearly independent set of generalized eigenvectors $\underset{\sim}{V}$ of decending rank.
The number of vectors in the set is called the **length** of the chain.

Canonical Basis

A **canonical basis** for an $n\times n$ matrix $\mathbf{A}$ is a set of n linearly independent generalized eigenvectors composed entirely of chains. That is, if a generalized eigenvector of rank$m$ appears in the basis, so does the complete chain generated by that vector.
The simplest canonical bases are those consisting solely of chains of length one (linearly independent eigenvectors); these exist when eigenvalues are distinct.

1.10 Similarity

$\mathbf{A}$ is similar to $\mathbf{B}$ if there exists invertible matrix $\mathbf{S}$ such that $\mathbf{A}=\mathbf{S}^{-1}\mathbf{BS}$.

1. Similar matrices have the same characteristic equation and, therefore, the same eigenvalues and the same trace.
2. If $\underset{\sim}{X}$ is an eigenvector of $\mathbf{A}$ associated with eigenvalues λ and $\mathbf{A}=\mathbf{S}^{-1}\mathbf{BS}$, then $\underset{\sim}{Y}=\mathbf{S}\underset{\sim}{X}$ is an eigenvector of $\mathbf{B}$ associated with the same eigenvalue.

If $\mathbf{A}$ is similar to $\mathbf{B}$, then $\mathbf{B}$ is similar to $\mathbf{A}$.

1.11 Inner Products

Complex Conjugates of $\mathbf{A}$ is $\bar{\mathbf{A}}$.

1. $\bar{\bar{x}}=x$ and $\bar{\bar{\mathbf{A}}}=\mathbf{A}$.
2. x is real iff $\bar{x}=x$, and $\mathbf{A}$ is a real matrix iff $\bar{\mathbf{A}}=\mathbf{A}$.
3. $x+\bar{x}$ is a real scalar and $\mathbf{A}+\bar{\mathbf{A}}$ is a real matrix.
4. $\overline{xy}=\bar{x}\bar{y}$ and $\overline{\mathbf{AB}}=\bar{\mathbf{A}}\bar{\mathbf{B}}$.
5. $\overline{x+y}=\bar{x}+\bar{y}$ and $\overline{\mathbf{A}+\mathbf{B}}=\bar{\mathbf{A}}+\bar{\mathbf{B}}$.
6. $x\bar{x}=|x|^2$ is real and positive (except when $x=0$).

Inner Product

For nonsingular $n\times n$ $\mathbf{W}$, the **inner produce** of $n\times 1$ vectors $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ with respect to $\mathbf{W}$, denoted $\langle\underset{\sim}{X},\underset{\sim}{Y}\rangle_{\mathbf{W}}$, is the dot product $\langle\underset{\sim}{X},\underset{\sim}{Y}\rangle_{\mathbf{W}}=(\mathbf{W}\underset{\sim}{X})\cdot(\overline{\mathbf{W}\underset{\sim}{Y}})$.
If $\mathbf{W}=\mathbf{I}$, then $\langle\underset{\sim}{X},\underset{\sim}{Y}\rangle=\underset{\sim}{X}\cdot\bar{\underset{\sim}{Y}}$ is the Euclidean inner product.
If $\underset{\sim}{X}$,$\underset{\sim}{Y}$ are real, then the Euclidean inner product is the dot product.

Inner product properties:

1. $\langle\underset{\sim}{X},\underset{\sim}{X}\rangle_{\mathbf{W}}$ is real and positive if $\underset{\sim}{X}\neq 0$.
2. $\langle\underset{\sim}{X},\underset{\sim}{X}\rangle_{\mathbf{W}}=0$ iff $\underset{\sim}{X}=\underset{\sim}{0}$.
3. $\langle\underset{\sim}{X},\underset{\sim}{Y}\rangle_{\mathbf{W}}=\overline{\langle\underset{\sim}{Y},\underset{\sim}{X}\rangle}_{\mathbf{W}}$.
4. $\langle c\underset{\sim}{X},\underset{\sim}{Y}\rangle_{\mathbf{W}}=c\langle\underset{\sim}{X},\underset{\sim}{Y}\rangle_{\mathbf{W}}$ and $\langle\underset{\sim}{X},c\underset{\sim}{Y}\rangle_{\mathbf{W}}=\bar{c}\langle\underset{\sim}{X},\underset{\sim}{Y}\rangle_{\mathbf{W}}$ for any scalar c, real or complex.
5. $\langle\underset{\sim}{X}+\underset{\sim}{Y},\underset{\sim}{Z}\rangle_{\mathbf{W}}=\langle\underset{\sim}{X},\underset{\sim}{Z}\rangle_{\mathbf{W}}+\langle\underset{\sim}{Y},\underset{\sim}{Z}\rangle_{\mathbf{W}}$ and $\langle\underset{\sim}{X},\underset{\sim}{Y}+\underset{\sim}{Z}\rangle_{\mathbf{W}}=\langle\underset{\sim}{X},\underset{\sim}{Y}\rangle_{\mathbf{W}}+\langle\underset{\sim}{X},\underset{\sim}{Z}\rangle_{\mathbf{W}}$.
6. Schwarz ineqality: $|\langle\underset{\sim}{X},\underset{\sim}{Y}\rangle_{\mathbf{W}}|^2\leq\langle\underset{\sim}{X},\underset{\sim}{X}\rangle_{\mathbf{W}}\langle\underset{\sim}{Y},\underset{\sim}{Y}\rangle_{\mathbf{W}}$.

Orthogonality

If the inner product of two vectors is zero they are **orthogonal**.
The inner product depends on $\mathbf{W}$.
A set of vectors is orthogonal when each vector is orthogonal to every other vector; such a set is linearly independent.
Gram-Schmidt orthogonalization process modifies a set of linearly independent vectors $\underset{\sim}{X}_i$ to an orthogonal set on nonzero vectors Q_i with respect to a specified inner product $\mathbf{W}$.

1.12 Norms

A **Vector Norm** is a measure of the length or magnitude of a vector, denoted $\|\underset{\sim}{X}\|$ and satisfying four conditions,

1. $\|\underset{\sim}{X}\|\geq 0$.
2. $\|\underset{\sim}{X}\|=0$ iff $\underset{\sim}{X}=\underset{\sim}{0}$.
3. $\|c\underset{\sim}{X}\|=|c|\;\|\underset{\sim}{X}\|$ for any scaler c.
4. Triangle inequality: $\|\underset{\sim}{X}+\underset{\sim}{Y}\|\leq\|\underset{\sim}{X}\|+\|\underset{\sim}{Y}\|$.

There are several common norms for measuring vector magnitude.

1. Inner-product-generated norm: $\|\underset{\sim}{X}\|_{\mathbf{W}}=\sqrt{\langle\underset{\sim}{X},\underset{\sim}{X}\rangle_{\mathbf{W}}}$.
2. Euclidean (or ℓ_2) norm: $\|\underset{\sim}{X}\|_2=\sqrt{\langle\underset{\sim}{X},\underset{\sim}{X}\rangle_{\mathbf{I}}}=\sqrt{\sum_{i=1}^n|x_i|^2}$.
3. ℓ_1 norm: $\|\underset{\sim}{X}\|_1=\sum_{i=1}^n|x_i|$.
4. ℓ_∞ norm: $\|\underset{\sim}{X}\|_\infty=\max_i(|x_i|)$.
5. ℓ_p norm $(p\geq 1)$: $\|\underset{\sim}{X}\|_p=(\sum_{i=1}^n|x_i|^p)^{1/p}$.

Normalized vectors and distance

A **unit vector** has norm equal to 1.
A nonzero vector is **normalized** when multiplied by the reciprocal of its norm, making it a unit vector.
A set of vectors is **orthnormal** if it is an orthogonal set of unit vectors.
The **distance** between two vectors $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ is $\|\underset{\sim}{X} - \underset{\sim}{Y}\|$.

Matrix norms

The norm of a matrix $\mathbf{A}$, $\|\mathbf{A}\|$ is a real-valued function satisfying five properties:

1. $\|\mathbf{A}\| \geq 0$.
2. $\|\mathbf{A}\| = \mathbf{0}$ iff $\underset{\sim}{X} = \underset{\sim}{0}$.
3. $\|c\mathbf{A}\| = |c| \ \|\mathbf{A}\|$ for any scaler c.
4. Triangle inequality: $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$.
5. Consistency condition: $\|\mathbf{AB}\| \leq \|\mathbf{A}\|\|\mathbf{B}\|$.

Because of the additional consistency condition, not all vector norms can be extended to be matrix norms. ℓ_1 and ℓ_2 can (sum of absolute values and sum of squares).
Euclidean (ℓ_2 or Frobenius) norm $\|\mathbf{A}\|_F = (\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2)^{1/2}$.

Induced norms

Over all vectors $\underset{\sim}{X}$ with norm equal to unity, a vector norm induces (or generates) **matrix norm** $\|\mathbf{A}\| = \max_{\|\mathbf{X}\|=1}(\|\mathbf{A}\underset{\sim}{X}\|)$ or the maximum element from the norm of the product $\mathbf{A}\underset{\sim}{X}$.

1. ℓ_1 is the largest column sum of absolute value.
2. ℓ_∞ is the largest row sum of absolute value.
3. ℓ_2 (spectral norm) is $\max \sqrt{\lambda}$, where λ is an eigenvalue of $\bar{\mathbf{A}}^T\mathbf{A}$.

Compatibility

A vector norm is **compatible** with a matrix norm if $\|\mathbf{A}\underset{\sim}{Y} \leq \|\mathbf{A}\|\|\underset{\sim}{Y}\|$, $\forall n \times n\mathbf{A}, n \times 1\underset{\sim}{Y}$.
Induced norms are always compatible and there always exists a $\underset{\sim}{Y}$ which gives equality.
Frobenius norm is compatible with Euclidean even through former not induced by latter.

Spectral Radius

The **spectral radius** os square matrix $\mathbf{A}$, $\sigma(\mathbf{A})$ is the largest absolute value of any eigenvalue of $\mathbf{A}$, that is, $\sigma(\mathbf{A}) = \max_i(\lambda_i)$.
For any matrix norm, $\sigma(\mathbf{A}) \leq \|\mathbf{A}\|$, (providing bounds on eigenvalues).
Equivelently, $\sigma(\mathbf{A}) = \lim_{m\to\infty} \|\mathbf{A}^m\|^{1/m}$.
$\sigma(\mathbf{A}^T) = \sigma(\mathbf{A})$.

Other

$\|\mathbf{I}\| = 1$.
Pythagorean theorem for inner-product-generated vector norm:
If $\langle \underset{\sim}{X}, \underset{\sim}{Y} \rangle_\mathbf{W} = 0$, then $\|\underset{\sim}{X} + \underset{\sim}{Y}\|^2_\mathbf{W} = \|\underset{\sim}{X}\|^2_\mathbf{W} + \|\underset{\sim}{Y}\|^2_\mathbf{W}$.
The **condition number** of square $\mathbf{A}$with respect to a matrix norm is $c(\mathbf{A}) = \begin{cases} \|\mathbf{A}\|\|\mathbf{A}^{-1}\| & \text{, if } \mathbf{A} \text{ nonsingular} \\ \infty & \text{, if } \mathbf{A} \text{ singular} \end{cases}$
Condition number of $\mathbf{I}$,$c(\mathbf{I}) = 1$ for all matrix norms.
$c(\mathbf{A}) \geq 1$.

1.13 Hermitian Matrices

Normal Matrices

The **Hermitian transpose** of $\mathbf{A}$, $\mathbf{A}^H$, is the complex conjugate transpose of $\mathbf{A}$, $\mathbf{A}^H = \bar{\mathbf{A}}^T$.
A matrix is **normal** if $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$.

1. Every normal matrix is similar to a diagonal matrix.
2. Every normal matrix possesses a canonical basis of eigenvectors which can be arranged to form an orthonormal set.

Hermitian Matrices

A matrix is hermitian if it equals its own Hermitian transpose, $\mathbf{A}^H = \mathbf{A}$.
The sum of Hermitian matricies is Hermitian, multiply one by a real scaler and it remains one, powers are Hermitian.
Hermitian matrix is normal because $\mathbf{A}\mathbf{A}^H = \mathbf{A}\mathbf{A} = \mathbf{A}^H\mathbf{A}$.

3. Has real eigenvalues.
4. —
5. $\mathbf{A}$ is Hermitian iff $\langle \mathbf{A}\underset{\sim}{X}, \underset{\sim}{X} \rangle$ is real for all (real and complex) vectors $\underset{\sim}{X}$.

Real Symmetric Matrices

$\mathbf{A}$ is **symmetric** if it equals its own transpose, that is $\mathbf{A}^T = \mathbf{A}$.
A **real symmetric** matrix is hermitian, and therefore, normal.

6. The eigenvalues of a real symmetric matrix can be chosen to be real

The Adjoint

The **adjoint** of an $n \times m$ matrix $\mathbf{A}$ is an $m \times n$ matrix $\mathbf{A}^*$ having the property that $\langle \underset{\sim}{X}, \mathbf{A}\underset{\sim}{Y} \rangle_\mathbf{W} = \langle \mathbf{A}^*\underset{\sim}{X}, \underset{\sim}{Y} \rangle_\mathbf{W}$ for all m-dim $\underset{\sim}{Y}$ and n-dim $\underset{\sim}{X}$.
The adjoint always exists and is $\mathbf{A}^* = (\mathbf{W}^H\mathbf{W})^{-1}\mathbf{A}^H(\mathbf{W}^H\mathbf{W})$.
When $\mathbf{W} = \mathbf{I}$, $A^* = A^H$.
$\mathbf{A}$ is **self-adjoint** if $A^* = A$.

Adjoint Identities

1. $(\mathbf{A}^*)^* = \mathbf{A}$.
2. $(\mathbf{A}+\mathbf{B})^* = \mathbf{A}^* + \mathbf{B}^*$.
3. $(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*$.
4. $(c\mathbf{A})^* = \bar{c}\mathbf{A}^*$ for any scalar c.

If $\underset{\sim}{X}$ is an eigenvector of a normal matrix $\mathbf{A}$ corresponding to eigenvlaue λ, then $\underset{\sim}{X}$ is the eigenvector of $\mathbf{A}^H$ corresponding to $\bar{\lambda}$.

Eigenvectors corresponding to distinct eigenvalues of a normal matrix are orthogonal with respect to the Euclidean inner product.

If $\mathbf{A}$ is Hermitian, then $\mathbf{A} - c\mathbf{I}$ is Hermitian for any real scalar c.

A Hermitian matrix is similar to a diagonal matrix.

A matrix is **skew-Hermitian** if $\mathbf{A} = -\mathbf{A}^H$, and

- $\mathbf{A}$ is normal.
- $i\mathbf{A}$ is Hermitian.
- $\langle \mathbf{A}\underset{\sim}{X}, \underset{\sim}{X}\rangle$ is pure imaginary for every $\underset{\sim}{X}$.
- then every eigenvalue of $\mathbf{A}$ is pure imaginary.

A matrix is **skew-symmetric** if $\mathbf{A} = -\mathbf{A}^T$.

- If real skew-symmetric, then skew-Hermitian.

1.14 Positive Definite Matrices

Definite Matrices

An $n \times n$ Hermitian matrix $\mathbf{A}$ is **positive definite (p.d.)** if $\langle \mathbf{A}\underset{\sim}{X}, \underset{\sim}{X}\rangle = 0$ for all non-zero $\underset{\sim}{X}$.

And $\mathbf{A}$ is **positive semi-definite (p.s.d.)** if $\langle \mathbf{A}\underset{\sim}{X}, \underset{\sim}{X}\rangle \geq 0$.

If inequalities are reversed, then **negative definite (n.d.)** and **negative semi-definite (p.s.d.)**.

The sum of two matrices of the same type are still that type.

p.d. (n.d.) matrices are invertible, and their inverses are of the same definite type.

Tests for Positive Definiteness

Any of three tests stipulate necessary and sufficient conditions for $n \times n$ Hermitian matrix $\mathbf{A}$ to be p.d.

If pass then p.d., if fail then not p.d.

1. $\mathbf{A}$ p.d. iff it can be reduced to upper triangular form using only elementary row operations E3 and the diagonal elements of the resulting matrix (the pivots) are all positive.
2. $\mathbf{A}$ p.d. iff all principal minors are positive. A **principal minor** is the determinant of a submatrix of $\mathbf{A}$ including first element.
3. $\mathbf{A}$ p.d. iff all eigenvalues are positive.

Any of three tests stipulate necessary conditions only for $n \times n$ Hermitian matrix $\mathbf{A}$ to be p.d. If fail not p.d., but no conclusion from passing.

4. Diagonal elements of $\mathbf{A}$ are positive.
5. Largest element of $\mathbf{A}$ in absolute value on diagonal.
6. $a_{ii}a_{jj} > |a_{ij}|^2$, $i \neq j$.

To test n.d., apply to $-\mathbf{A}$.

Square roots of matrices

The **square root** of $\mathbf{A}$ is $\mathbf{A}^{1/2}$ such that $\mathbf{A} = \mathbf{A}^{1/2}\mathbf{A}^{1/2}$.

If $\mathbf{A}$and $\mathbf{A}^{1/2}$ are both p.d. or p.s.d., then $\mathbf{A}^{1/2}$ is unique and square root runction is well defined.

Cholesky Decomposition

P.d. matrix $\mathbf{A}$ can be factorized into $\mathbf{A} = \mathbf{L}\mathbf{L}^H$ where $\mathbf{L}$ is a lower triangular matrix with positive values on the diagonal.

If $\mathbf{A}$ is p.d. and $\mathbf{C}$ is nonsingular, then $\mathbf{B} = \mathbf{C}^H\mathbf{A}\mathbf{C}$ is p.d. If $\mathbf{A}$ is Hermitian, $e^{\mathbf{A}t}$ is p.d.

1.15 Unitary Transformations

Unitary Matrices

A matrix is **unitary** if its inverse equals its Hermitian transpose, $\mathbf{U}^{-1} = \mathbf{U}^H = \bar{\mathbf{U}}^T$.

Properties:

1. Unitary iff columns (or rows) form an orthogonal set.
2. A product of unitary matrices is unitary.
3. If $\mathbf{U}$ is unitary, $\langle \mathbf{U}\underset{\sim}{X}, \mathbf{U}\underset{\sim}{Y}\rangle = \langle \underset{\sim}{X}, \underset{\sim}{Y}\rangle$.
4. All eigenvalues have absolute value 1.
5. Determinant has absolute value 1.

An orthogonal matrix with all real elements is unitary.

If $\mathbf{P}$ is orthogonal, $\mathbf{P}^{-1} = \mathbf{P}^T$.

Schur Decomposition

Every square matrix is similar to an upper triangular matrix, and a unitary matrix can be chosen to produce the transformation.

This transformation $\mathbf{U}^H\mathbf{A}\mathbf{U} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}$ is triangular is the Schur decomposition.

Elementary Reflectors

An elementary reflector (or Householder transformation) associated with real vector $\underset{\sim}{V}$ yields matrix $\mathbf{R}$ which is real symmetric, orthogonal and whose square is the identity matrix.

$\mathbf{R} = \mathbf{I} - \frac{2\underset{\sim}{V}\underset{\sim}{V}^T}{\|\underset{\sim}{V}\|_2^2}$, (using ℓ_2 norm (SS)).

$\|\mathbf{U}\underset{\sim}{X}\|_2 = \|\underset{\sim}{X}\|_2$, thus unitary transformation preserves Euclidean length.

Angle between two vectors $\theta = \arccos\left\{\frac{\langle \underset{\sim}{X}, \underset{\sim}{Y}\rangle}{\|\underset{\sim}{X}\|_2\|\underset{\sim}{Y}\|_2}\right\}$, and unitary transformation preserves angles.

Rotation matrix

The **rotation matrix** $\mathbf{R}_{p,q}(\theta)$ if

- diagonal: (p,p) and (q,q) elements are $\cos(\theta)$ for $p \neq q$ and all other diagonals are 1.
- off-diagonal: (p,q) elements are $\sin(\theta)$, (q,p) elements are $-\sin(\theta)$.
- others: are 0.

A rotational matrix is orthogonal.

1.16 Quadratic Forms and Congruence

A **quadratic form** in $\underset{\sim}{X}$ is a polynomial of type $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$ with real-valued coefficients a_{ij} or $\underset{\sim}{X}^T\mathbf{A}\underset{\sim}{X}$ or $\underset{\sim}{X}^T\frac{\mathbf{A}+\mathbf{A}^T}{2}\underset{\sim}{X}$, since $\frac{\mathbf{A}+\mathbf{A}^T}{2}$ is symmetric.

A **complex quadratic form** is when $\underset{\sim}{X}$ is Hermitian.

Both are equivalent to inner-product $\langle \mathbf{A}\underset{\sim}{X}, \underset{\sim}{X}\rangle$, which is real when $\mathbf{A}$ is Hermitian.

If this inner product is positive (neg, etc.) then quadratic form is p.d. (n.d., etc.).

A quadratic form has **diagonal form** if it contains no cross-product terms.

1.17 Nonnegative Matrices

$\mathbf{A}$ is nonnegative if all elements are real and nonnegative.

$\mathbf{A}$ is greater than $\mathbf{B}$ if $\mathbf{A}-\mathbf{B}$ is positive.

Properties of Spectral Radii of nonnegative square matrices

1. If $\mathbf{0} \le \mathbf{A} \le \mathbf{B}$ then $\sigma(\mathbf{A}) \le \sigma(\mathbf{B})$.
2. If $\mathbf{A} \ge \mathbf{0}$ and row (or column) sums of $\mathbf{A}$ are constant k, then $\sigma(\mathbf{A}) = k$.
3. If m is minimum row (or column) sum, and M is maximum, and $\mathbf{A} \ge 0$, then $m \le \sigma(\mathbf{A}) \le M$.
4. A nonnegative square matrix has eigenvalue equal its spectral radius, and there exists left and right eigenvectors with nonnegative components corresponding to this eigenvalue.
5. (Perron's theorem) A positive square matrix has its largest eigenvalue of multiplicity one equal to its spectral radius.

Irreducible Matrices

A **permutation matrix** is a matrix obtained from an identity matrix by any rearrangement of its row.

$\mathbf{A}$ is **reducible** if there exists a permutation matrix $\mathbf{P}$ such that $\mathbf{PAP}^T = \left[\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{0} & \mathbf{A}_{22} \end{array}\right]$, (an upper-triangular partitioned matrix).

If no such permutation matrix exists, then $\mathbf{A}$ is **irreducible**.

1. Positive matrices are irreducible.
2. An $n \times n$ matrix $\mathbf{A}$ is irreduible iff $(\mathbf{I}+\mathbf{A})^{n-1} \ge \mathbf{0}$.
3. (Perron-Frobenius theorem) A nonnegative, irreducible matrix $\mathbf{A}$ has an eigenvalue of multiplicity one equal to its spectral radius and there are right or left eigenvectors corresponding to this eigenvalue with only positive components. If $\mathbf{A}$ has k eigenvalues in absolute value equal to its spectral radius, then they are of the form $w_i\sigma(\mathbf{A})$, where $w_1, \ldots, w_k$ are the k distinct roots of unity.

Primitive Matrices

A nonnegative matrix is **primitive** if it is irreducible and has only one eigenvalue with absolute value equal its spectral radius.

A nonnegative matrix is **regular** if one of its powers is positive.

A nonnegative matrix is primitive iff it is regular.

If $\mathbf{A}$ is a nonnegative primative matrix, then the limit $\mathbf{L} = \lim_{m\to\infty}\left\{\frac{\mathbf{A}}{\sigma(\mathbf{A})}\right\}^m$ exists and is positive. Furthermore, if $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ are the left and right positive eigenvectors of $\mathbf{A}$ corresponding to eigenvalue equal to $\sigma(\mathbf{A})$ and scaled so that $\underset{\sim}{Y}\underset{\sim}{X} = 1$, then $\mathbf{L} = \underset{\sim}{X}\underset{\sim}{Y}$.

Stochastic Matrices

A nonnegative matrix is **stocastic** if all its row (or column) sums equal 1.

It is **doubly stochastic** if both row and columns equal 1.

The spectral radius of a stochastic matrix is 1.

The right (or left) eigenvector corresponding to $\lambda = 1$ has all its components equal.

A stochastic matrix is **ergodic** if the only eigenvalue of absolute value 1 is 1 itself, and if $\lambda = 1$ has multiplicity k, then there exists k lineraly independent eigenvectors corresponding to it.

If $\mathbf{P}$ is ergodic, then $\lim_{m\to\infty}\mathbf{P}^m = \mathbf{L}$ exists.

All rows of $\mathbf{L}$ are identical and is the unique left eigenvector corresponding to $\lambda = 1$ and having the sum of its components equal to unity.

The limiting matrix is not as simple for an ergodic matrix that is not primitive.

A canonical basis for such a matrix consists solely of eigenvectors.

If the multiplicity of $\lambda = 1$ is denoted by k, and if the k linearly independent right eigenvectors corresponding to this eigenvalue are placed into the first k columns of modal matrix $\mathbf{M}$, then $\mathbf{L} = \mathbf{MDM}^{-1}$, where $\mathbf{D}$ is diagonal matrix having first k diagonal elements equal to 1, and all others zero.

Finite Marcov Chains

An N**-state Marcov chain** consists of a set of objects and a finite set of N different states such that

1. At any given time each object is in one of the N states, and
2. The probability that an object will move from one state to another (or remain in the same state) in one time period depends only on the beginning and ending states.

The $N\times N$ matrix $\mathbf{P} = [p_{ij}]$, where p_{ij} denotes the probability of transitioning from state i to state j in one time period, is stochastic.

The (i,j) element of the mth power of $\mathbf{P}$ represents the probability an object will move from state i to state j in m time periods.

Denote the proportion of objects in state i at the end of the mth time period as $x_i^{(m)}$, and define $\underset{\sim}{X}^{(m)} = [x_1^{(m)}, \ldots, x_N^{(m)}]^T$ to be the **distribution vector** for the end of the mth time period.

Then $\underset{\sim}{X}^{(0)}$ represents the proportion of objects in each state at the beginning of the process.

$\underset{\sim}{X}^{(m)} = \underset{\sim}{X}^{(0)}\mathbf{P}^m$.

If $\mathbf{P}$ is primitive, then $\underset{\sim}{X}^{(\infty)} = \lim_{m\to\infty}\underset{\sim}{X}^{(m)} = \underset{\sim}{X}^{(0)}\mathbf{L}$ which is the left eigenvector of $\mathbf{P}$ corresponding to $\lambda = 1$ and having the sum of its components equal to 1.

The ith component of $\underset{\sim}{X}^{(\infty)}$ represents the approximate proportion of objects in state i after a large number of time periods, and its limiting distribution is independent of $\underset{\sim}{X}^{(0)}$.

If **P** is ergodic but not primitive, the limit $\lim_{m\to\infty} \underset{\sim}{X}^{(m)} = \underset{\sim}{X}^{(0)}\mathbf{L}$ still may be used to obtain the limiting state distribution, but it will depend on $\underset{\sim}{X}^{(0)}$.

If **A**, **B**are stochastic matrices, then $\mathbf{C} = \mathbf{AB}$ is stocastic.

1.18 Patterned Matrices

Circulant Matrices

A **circulant matrix** is square with each row being the same as the previous but shifted one column to the right, with the last element becoming the first.

Diagonals have identical elements.

1. For $n \times n$ circulant matrix $\mathbf{A}$, its eigenvalues are $\lambda_i = a_1 + a_2 r_i + a_3 r_i^2 + \cdots + a_n r_i^{n-1}$, $(i = 1, \ldots, n)$ where $[a_1, a_2, \ldots, a_n]$ is the first row of $\mathbf{A}$ and r_i is one of the n distinct solutions of $r^n = 1$. The corresponding eigenvectors are $\underset{\sim}{X}_i = [1, r_i, r_i^2, \ldots, r_i^{n-1}]^T$.
2. For **A**,**B**circulant, $a\mathbf{A} + b\mathbf{B}$ is circulant.
3. The product of circulant matricies is circulant and multiplication is commutative.
4. For a nonsingular circulant matrix, its inverse is circulant.