Review of Linear Algebra – E186 Handout

Vectors and Their Inner Products

Let X and Y be two vectors:

$$X = [x_1, \cdots, x_n]^T$$

and

$$Y = [y_1, \cdots, y_n]^T$$

Their *inner product* is defined as

$$(X, Y) \triangleq X^{*T} Y = \sum_{k=1}^{n} x_k^* y_k$$

where T and $*$ represent transpose and complex conjugate, respectively.

The *norm* (magnitude, length) of a vector X is defined as

$$\|X\| \triangleq (X, X)^{1/2} = \sqrt{\sum_{k=1}^{n} |x_k|^2}$$

where $|x|$ represents the absolute value of x (real or complex). X is normalized if $\|X\| = 1$.

Two vectors X and Y are *orthogonal* to each other iff their inner product is zero. For normalized orthogonal vectors, we have

$$(X, Y) = \delta_{XY} \triangleq \begin{cases} 1 & if\ X = Y \\ 0 & if\ X \neq Y \end{cases}$$

Rank, Trace, Determinant, Transpose and Inverse of a Matrix

Let A be an $n \times n$ square matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdot & a_{nn} \end{bmatrix}_{n \times n}$$

where

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \cdots \\ a_{nj} \end{bmatrix}$$

is the jth column vector and

$$[a_{i1}\ a_{i2}\ \cdots\ a_{in}]$$

is the ith row vector.

The n rows span the *row space* of A and the n columns span the *column space* of A. The dimensions of these two spaces are the same and called the *rank* of A:

$$R = rank(A) \leq N$$

The *determinant* of A is denoted by

$$det(A) = |A|$$

and we have

$$|AB| = |A|\ |B|$$

$rank(A) < N$ iff $det(A) = 0$.

The *trace* of A is defined as the sum of its diagonal elements:

$$tr(A) = \sum_{i=1}^{n} a_{ii}$$

The *transpose* of a matrix A, denoted by A^T, is obtained by switching the positions of elements a_{ij} and a_{ji} for all $i, j \in \{1, \cdots, n\}$. In other words, the ith column of A becomes the ith row of A^T, or equivalently, the ith row of A becomes the ith column of A^T:

$$A^T = [A_1 \cdots A_n]^T = \begin{bmatrix} A_1^T \\ \cdot\cdot \\ \cdot\cdot \\ A_n^T \end{bmatrix}$$

where vector A_i is the ith column of A and its transpose A_i^T is the ith row of A^T.

For any two matrices A and B, we have

$$(AB)^T = B^T A^T$$

If $AB = BA = I$, where I is an identity matrix:

$$I = diag[1, \cdots, 1] = \begin{bmatrix} 1 & 0 & \cdot & 0 \\ 0 & 1 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & 1 \end{bmatrix}$$

then $B = A^{-1}$ is the *inverse* of A. A^{-1} exists iff $det(A) \neq 0$, i.e., $rank(A) = N$.

For any two matrices A and B, we have

$$(AB)^{-1} = B^{-1} A^{-1}$$

and

$$(A^{-1})^T = (A^T)^{-1}$$

Hermitian Matrix and Unitary Matrix

A is a *Hermitian matrix* iff $A^{*T} = A$. When a Hermitian matrix A is real ($A^* = A$), it becomes a *symmetric matrix*, $A^T = A$.

A is a *unitary matrix* iff $A^{*T}A = I$, i.e.,$A^{*T} = A^{-1}$. When a unitary matrix A is real ($A^* = A$), it becomes an *orthogonal matrix*, $A^T = A^{-1}$.

The columns (or rows) of a unitary matrix A are *orthonormal*, i.e. they are both orthogonal and normalized, i.e.,

$$(A_i, A_j) = \delta_{ij} \triangleq \begin{cases} 1 & if\ i = j \\ 0 & if\ i \neq j \end{cases}$$

where A_i and A_j represent the ith and jth columns of A, respectively.

As we will see later, any Hermitian matrix A can be converted to a diagonal matrix Λ (or diagonalized) by a particular unitary matrix Φ:

$$\Phi^{*T} A \Phi = \Lambda$$

where Λ is a diagonal matrix, i.e., all its off diagonal elements are 0.

Unitary Transforms

For any given unitary matrix $A = [A_1 \; A_2 \; \cdots \; A_n]$ (where vector A_i is the ith colum of the matrix), a *unitary transform* of a vector $X = [x_1, x_2, \cdots, x_n]^T$ can be defined as

$$\left\{ \begin{array}{l} Y = A^{*T} X = \left[\begin{array}{c} A_1^T \\ ... \\ A_n^T \end{array} \right] X \\ X = AY = [A_1, A_2, \; \cdots, \; A_n] \; Y \end{array} \right.$$

where $Y = [y_1, y_2, \cdots, y_n]^T$ is another vector.

The first equation of the unitary transform is the forward transform and the ith component of Y can be written as:

$$y_i = A_i^{*T} X = (A_i, X)$$

y_i is the inner product of X and the ith column vector A_i of A, i.e., the projection of X on the ith vector A_i.

The second equation is the inverse transform and can be written as

$$X = \sum_{i=1}^{n} y_i A_i$$

i.e., vector X is represented as a linear combination (weighted sum) of the n column vectors $A_i, A_2, \cdots, A_n$ of the transform matrix A. In other words, X is represented as a vector (or a point) in the n-dimensional space spanned by the n orthonormal column vectors $A_1, A_2, \cdots, A_n$. Each of the n coordinates $(y_1, y_2, \cdots, y_n)$ of this vector is its projection on the direction specified by the corresponding column vector of A.

Specially when $A = I$, we have

$$X = \sum_{i=1}^{n} y_i A_i = \sum_{i=1}^{n} x_i I_i$$

where $I_i = [0, \cdots, 0, 1, 0, \cdots, 0]^T$ is the ith column of the identity matrix I with the ith element equal 1 and all other 0.

Unitary transform does not change a vector's norm:

$$||Y||^2 = Y^{*T}Y = (A^{*T}X)^{*T}(A^{*T}X) = X^{*T}AA^{*T}X = X^{*T}X = ||X||^2$$

as $AA^{*T} = I$. In other words, the length of a vector is always the same in different coordinate systems.

The geometric interpretation of any unitary transform $Y = AX$ is to rotate a vector about the origin $[0, \cdots, 0]^T$ (rotation does not change the vector's length), or equivalently, to represent the same vector X by the coordinates Y in a different coordinate system.

If X is interpreted as a signal, then $||X||^2$ can be interpreted as the total energy or information contained in the signal which is preserved by during any unitary transform. However, some other features of the signal may be changed, e.g., the signal may be decorrelated after the transform, which is desirable in many applications.

Eigenvalues and Eigenvectors

For any matrix A, if there exist a vector ϕ and a value λ such that

$$A\phi = \lambda\phi$$

then λ and ϕ are called the *eigenvalue* and *eigenvector* of matrix A, respectively. To obtain λ, rewrite the above equation as

$$(\lambda I - A)\phi = 0$$

which is a homogeneous equation system. To find its non-zero solution for ϕ, we require

$$|\lambda I - A| = 0$$

Solving this nth order equation of λ, we get n eigenvalues $\{\lambda_1, \cdots, \lambda_n\}$. Substituting each λ_i back into the equation system, we get the corresponding eigenvector ϕ_i. We now have

$$A[\phi_1, \cdots, \phi_n] = [\lambda_1\phi_1, \cdots, \lambda_n\phi_n] = [\phi_1, \cdots, \phi_n] \begin{bmatrix} \lambda_1 & 0 & \cdot & 0 \\ 0 & \lambda_2 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \lambda_n \end{bmatrix}$$

or in a more compact form,

$$A\Phi = \Phi\Lambda$$

or

$$\Phi^{-1}A\Phi = \Lambda$$

where

$$\Phi = [\phi_1, \cdots, \phi_n]$$

and

$$\Lambda = diag[\lambda_1, \cdots, \lambda_n]$$

The trace and determinant of A can be obtained from its eigenvalues

$$tr(A) = \sum_{k=1}^{n} \lambda_k$$

and

$$det(A) = \prod_{k=1}^{n} \lambda_k$$

A^T has the same eigenvalues and eigenvectors as A.

A^m has the same eigenvectors as A, but its eigenvalues are $\{\lambda_1^m, \cdots, \lambda_n^m\}$, where m is a positive integer.

This is also true for $m = -1$, i.e., the eigenvalues of A^{-1} are $\{1/\lambda_1, \cdots, 1/\lambda_n\}$.

If A is Hermitian (symmetric if A is real), all the λ_i's are real and all eigenvectors ϕ_i's are orthogonal:

$$(\phi_i, \phi_j) = \delta_{ij}$$

If all ϕ_i's are normalized, matrix Φ is unitary (orthogonal if A is real):

$$\Phi^{-1} = \Phi^{*T}$$

and we have

$$\Phi^{-1} A \Phi = \Phi^{*T} A \Phi = \Lambda$$

Positive Definite Matrix

A real symmetric matrix A is *positive definite*, denoted by $A > 0$, iff the quadratic form X^TAX is greater than zero:

$$X^TAX > 0$$

for any $X = [x_1, \cdots, x_n]^T$ (x_i's are not all zero).

$A > 0$ iff all its eigenvalues are greater than zero:

$$\lambda_i > 0, \quad i = (1, \cdots, n)$$

As the eigenvalues of A^{-1} are $1/\lambda_i,\ i = (1, \cdots, n)$, we have $A > 0$ iff $A^{-1} > 0$.

Vector Differentiation

A vector differentiation operator is defined as

$$\frac{d}{dX} \triangleq [\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_n}]^T$$

which can be applied to any scalar function $f(X)$ to find its derivative with respect to X:

$$\frac{d}{dX} f(X) = [\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n}]^T$$

Vector differentiation has the following properties:

$$\frac{d}{dX}(B^T X) = \frac{d}{dX}(X^T B) = B$$

$$\frac{d}{dX}(X^T X) = 2X$$

$$\frac{d}{dX}(X^T A X) = 2AX \quad (if\ A^T = A)$$

To prove the third one, consider the kth element of the vector:

$$\frac{\partial}{\partial x_k}(X^T A X) = \frac{\partial}{\partial x_k} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j = \sum_{i=1}^{n} a_{ik} x_i + \sum_{j=1}^{n} a_{kj} x_j = 2 \sum_{i=1}^{n} a_{ik} x_i$$

for $(k = 1, \cdots, n)$.

Note that here we have used the assumption that $a_{ik} = a_{ki}$, i.e., $A^T = A$. Putting all n elements in vector form, we have the above.

When $A = I$, we have

$$\frac{d}{dX}(X^T X) = 2X$$

You can compare these results with the familiar derivatives in the scalar case:

$$\frac{d}{dx}(ax^2) = 2ax$$