1.2 MATRICES

Definition of Matrices

Let A be the linear mapping from a vector x to another vector y. Then A is represented by a rectangular array of numbers:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

that is called $m \times n$-matrix.

Transpose of a Matrix A:

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

Conjugate Transpose of a Matrix A:

$$A^* = \begin{bmatrix} \bar{a}_{11} & \bar{a}_{21} & \cdots & \bar{a}_{m1} \\ \bar{a}_{12} & \bar{a}_{22} & \cdots & \bar{a}_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ \bar{a}_{1n} & \bar{a}_{2n} & \cdots & \bar{a}_{mn} \end{bmatrix}$$

Notice that $A^T = A^*$ for real matrices.

Basic Operation of Matrices

a: a scalar, A, B: matrices

Addition:

$$A + B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

Scalar Multiplication:

$$aA = a \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} aa_{11} & aa_{12} & \cdots & aa_{1n} \\ aa_{21} & aa_{22} & \cdots & aa_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ aa_{m1} & aa_{m2} & \cdots & aa_{mn} \end{bmatrix}$$

Basic Operation of Matrices (Continued)

Matrix Multiplication:

$$AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1l} \\ b_{21} & b_{22} & \cdots & b_{2l} \\ \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nl} \end{bmatrix}$$

$$= \begin{bmatrix} \Sigma_{i=1}^{n} a_{1i}b_{i1} & \Sigma_{i=1}^{n} a_{1i}b_{i2} & \cdots & \Sigma_{i=1}^{n} a_{1i}b_{il} \\ \Sigma_{i=1}^{n} a_{2i}b_{i1} & \Sigma_{i=1}^{n} a_{2i}b_{i2} & \cdots & \Sigma_{i=1}^{n} a_{2i}b_{il} \\ \vdots & \vdots & \vdots & \vdots \\ \Sigma_{i=1}^{n} a_{mi}b_{i1} & \Sigma_{i=1}^{n} a_{mi}b_{i2} & \cdots & \Sigma_{i=1}^{n} a_{mi}b_{il} \end{bmatrix}$$

Inverse of Square Matrices

Inverse of an $n \times n$ matrix A is an $n \times n$ matrix such that

$$AA^{-1} = A^{-1}A = I$$

Theorem: An $n \times n$ matrix A has its inverse iff the columns of A are linearly independent.

Suppose A defines a linear transformation:

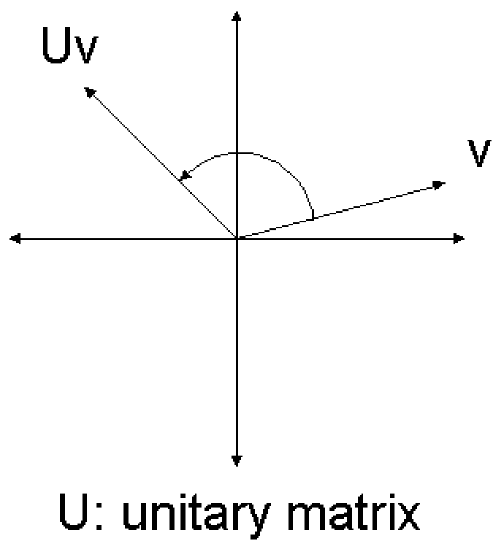
$$y = Ax$$

Then the inverse of A defines the inverse transformation:

$$x = A^{-1}y$$

Unitary Matrices

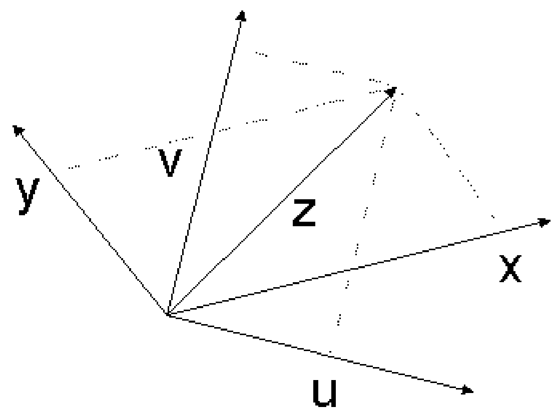
Matrix that rotates a vector without change of size is called unitary matrix



Properties of Unitary Matrix:

$$U^*U = I = UU^*$$

Coordinate Transformation



z: a vector, $\{u, v\}$, $\{x, y\}$: coordinate systems

$$z = \bar{w}_1 u + \bar{w}_2 v = w_1 x + w_2 y \quad \Rightarrow \quad [u\ v]\bar{w} = [x\ y]w$$

$$\Downarrow$$

$$\bar{w} = Tw, \quad T = [u\ v]^{-1}[x\ y]$$

In general, the representations of a vector in two different coordinate systems are related by an invertible matrix T:

$$\bar{w} = Tw$$

$$w = T^{-1}\bar{w}$$

Coordinate Transformation (Continued)

Representations of matrix in different coordinates:

Suppose $\alpha = T\bar{\alpha}$ and $\beta = T\bar{\beta}$. Then

$$\alpha = A\beta \quad \Rightarrow \quad T\bar{\alpha} = AT\bar{\beta} \quad \Rightarrow \quad \bar{\alpha} = T^{-1}AT\bar{\beta}$$

$$\Downarrow$$

$$\bar{\alpha} = \bar{A}\bar{\beta}$$

where

$$\bar{A} = T^{-1}AT$$

that is called the similarity transformation of A.

Eigenvalues and Eigenvectors

The eigenvalues of $n \times n$ matrix A are n roots of $det(\lambda I - A)$.

If λ is an eigenvalue of A, $\exists$ nonzero v such that

$$Av = \lambda v$$

where v is called eigenvector.

Eigenvalue Decomposition

Let $A \in \mathbf{R}^{n \times n}$. Suppose λ_i be eigenvalues of A such that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$$

Let

$$T = [v_1, v_2, \cdots, v_n] \in \mathbf{R}^{n \times n}$$

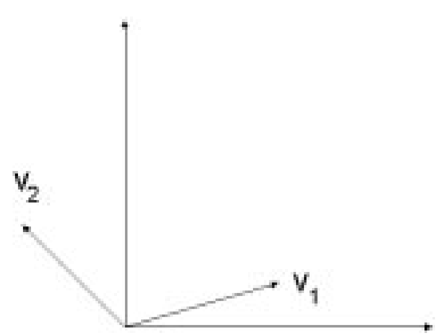
where v_i denotes eigenvector of A associated with λ_i. If A has n linearly independent eigenvectors,

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

where

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

Notice that Λ is simply the representation of A in the coordinate system consists of eigenvectors.



Symmetric Matrices

A matrix A is called symmetric if

$$A = A^T$$

Symmetric matrix is useful when we consider a quadratic form.

Indeed, given a matrix A,

$$x^T A x = x^T S x$$

where S is the symmetric matrix defined by

$$S = \frac{1}{2}(A + A^T)$$

Positive Definiteness: A symmetric matrix A is positive definite if

$$x^T A x > 0 \quad \forall x \neq 0,\ x \in \mathbf{C}^n$$

Positive Semi-Definiteness: A symmetric matrix A is positive semi-definite if

$$x^T A x \geq 0$$

Theorem: A symmetric matrix A is positive definite iff all the eigenvalues of A are positive.

Matrix Norms

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbf{C}^{m \times n}$$

p norms:

$$|||A|||_p = (\sum_{i,j} |a_{i,j}|^p)^{\frac{1}{p}} \qquad 1 \leq p < \infty$$

$$|||A|||_\infty = \max_{i,j} |a_{i,j}|$$

$||| \cdot |||_2$ is called Euclidean or Frobenius norm.

What is the difference between $\mathbf{C}^{m \times n}$ and $\mathbf{C}^{mn}$?

A matrix in $\mathbf{C}^{m \times n}$ defines a linear operator from $\mathbf{C}^n$ to $\mathbf{C}^m$; $y = Ax$.

Matrix Norms (Continued)

Induced (or operator) p norms:

$$\|A\|_p = \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p} = \max_{\|x\|=1} \|Ax\|_p \qquad 1 \leq p \leq \infty$$

$$\Downarrow$$

$$\|y\|_p = \|Ax\|_p \leq \|A\|_p \|x\|_p \qquad \forall x \in \mathbf{C}^n$$

Examples:

$p = 1$:

$$\|A\|_1 = \max_j \sum_{i=1}^{m} |a_{i,j}|$$

$p = 2$: spectral norm

$$\|A\|_2 = [\lambda_{max}(A^T A)]^{\frac{1}{2}}$$

$p = \infty$:

$$\|A\|_\infty = \max_i \sum_{j=1}^{m} |a_{i,j}|$$