Review on Matrices

- **Reading Assignments**

H. Anton and C. Rorres, *Elementary Linear Algebra* (Applications Version), 8th edition, John Wiley, 2000 (1.3-1.4, hard copy).

J. Pricipe et al., *Neural and Adaptive Systems: Fundamentals Through Simulations*, (Appendix A: Elements of Linear Algebra and Pattern Recognition, pp. 590-594, hard copy).

K. Kastleman, *Digital Image Processing*, Prentice Hall, (Appendix 3: Mathematical Background, hard copy).

F. Ham and I. Kostanic. *Principles of Neurocomputing for Science and Engineering*, Prentice Hall, (Appendix A: Mathematical Foundation for Neurocomputing, hard copy)

- **Other Books**

B. Kolman and D. Hill, *Introductory Linear Algebra with Applications*, 2nd edition, Prentice Hall, 2001.

L. Johnson, R. Riess, and J. Arnold, *Introduction to Linear Algebra*, 4th edition, Addison Wesley, 1998.

Review on Matrices

• Matrix addition/subtraction

- Matrices can be added or subtracted as long as they are of the same dimension.

$$C = A + B \text{ implies } c_{ij} = a_{ij} + b_{ij}$$

$$C = A - B \text{ implies } c_{ij} = a_{ij} - b_{ij}$$

• Matrix multiplication

$$\begin{bmatrix} a_{11} & a_{12} & . & a_{1n} \\ a_{21} & a_{22} & . & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & . & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & . & b_{1p} \\ b_{21} & b_{22} & . & b_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ b_{q1} & b_{q2} & . & b_{qp} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & . & c_{1p} \\ c_{21} & c_{22} & . & c_{2p} \\ \cdots & \cdots & c_{ij} & \cdots \\ c_{m1} & c_{m2} & . & c_{mp} \end{bmatrix}$$

# of columns of matrix A = # of rows of matrix B

(if A is m x n and B is q x p, then C will be m x p (**assume n=q**))

$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}$$

Example: $$\begin{bmatrix} 2 & 3 \\ 4 & 0 \\ 3 & -2 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 5 & -1 & 3 \end{bmatrix} = \begin{bmatrix} 17 & 1 & 9 \\ 4 & 8 & 0 \\ -7 & 8 & -6 \\ 10 & 9 & 3 \end{bmatrix}$$

• Properties of matrix multiplication

$$A(B + C) = AB + AC \text{ (distributive law)}$$

$$AB \neq BA$$

$$AI = IA = A, \text{ where } I = \begin{bmatrix} 1 & 0 & . & 0 \\ 0 & 1 & . & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & . & 1 \end{bmatrix}$$

• **Matrix transpose**

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

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

• **Symmetric Matrix** (matrix must be square)

- $A = A^T$ $(a_{ij} = a_{ji})$

Example: $\begin{bmatrix} 4 & 5 & -3 \\ 5 & 7 & 2 \\ -3 & 2 & 10 \end{bmatrix}$

• **Determinants** (matrix must be square)

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

$$det(A) = \sum_{j=1}^{m} (-1)^{j+k} a_{jk} det(A_{jk}), \text{ for any } k\text{: } 1 \le k \le m$$

$$det(AB) = det(A)det(B)$$

$$det(A+B) \ne det(A) + det(B)$$

$$\text{If } A = \begin{bmatrix} a_{11} & 0 & . & 0 \\ 0 & a_{22} & . & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & . & a_{nn} \end{bmatrix}, \text{ then } det(A) = \prod_{i=1}^{n} a_{ii}$$

• **Matrix inverse** (matrix must be square)

- The inverse A^{-1} of matrix A has the property: $AA^{-1}=A^{-1}A=I$

- A^{-1} exists only if $det(A) \neq 0$

singular: the inverse of A does not exist

ill-conditioned: A is nonsingular but close to being singular

- Some properties of the inverse:

$$det(A^{-1}) = \frac{1}{det(A)}$$

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

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

• **Pseudo-inverse**

- If A is not square (i.e., mxn), then its pseudo-inverse A^+ is given by:

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

- You can easily show that

$$A^+ A = I \quad \text{(provided that } (A^T A)^{-1} \text{ exists)}$$

• **Trace of a matrix** (matrix must be square)

$$tr(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

$$tr(A^T) = tr(A)$$

$$tr(A \pm B) = tr(A) \pm tr(B)$$

$$tr(AB) = tr(BA)$$

(in general, $tr(AB) \neq tr(A)tr(B)$)

• **Rank of a matrix**

- It is equal to the dimension of the largest square submatrix of A that has a non-zero determinant.

Example:

$$A = \begin{bmatrix} 4 & 5 & 2 & 14 \\ 3 & 9 & 6 & 21 \\ 8 & 10 & 7 & 28 \\ 1 & 2 & 9 & 5 \end{bmatrix} \text{has rank 3}$$

$$det(A) = 0, \text{ but } det(\begin{bmatrix} 4 & 5 & 2 \\ 3 & 9 & 6 \\ 8 & 10 & 7 \end{bmatrix}) = 63 \neq 0$$

- Alternatively, it is the maximum number of linearly independent columns or rows of A.

Example (cont'd):

$$1\begin{bmatrix} 4 \\ 3 \\ 8 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 5 \\ 9 \\ 10 \\ 2 \end{bmatrix} + 0\begin{bmatrix} 2 \\ 6 \\ 7 \\ 9 \end{bmatrix} - 1\begin{bmatrix} 14 \\ 21 \\ 28 \\ 5 \end{bmatrix} = 0$$

• **Matrix properties based on rank**

(1) If A is mxn, $rank(A) \leq \min m, n$

(2) If A is nxn, $rank(A) = n$ iff A is nonsingular (i.e., invertible).

(3) If A is nxn, $rank(A) = n$ iff $det(A) \neq 0$ (**full rank**).

(4) If A is nxn, $rank(A) < n$ iff A is singular

• Orthogonal/orthonormal matrices

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

- Consider the vectors formed by the rows (or columns) of matrix A:

$$u_1^T = [a_{11}\ a_{12} \cdots a_{1n}]$$
$$u_2^T = [a_{21}\ a_{22} \cdots a_{2n}]$$
$$...$$
$$u_m^T = [a_{m1}\ a_{m2} \cdots a_{mn}]$$

$$\text{or } A = \begin{bmatrix} u_1^T \\ u_2^T \\ ... \\ u_m^T \end{bmatrix} \text{(get used to this notation !)}$$

- Consider the following two properties:

(1) $u_k . u_k = 1$ or $||u_k|| = 1$, for every k

(2) $u_j . u_k = 0$, for every $j \neq k$ (u_j is perpendicular to u_k)

<u>A is orthonormal</u> if both (1) and (2) are satisfied

<u>A is orthogonal</u> if only (2) is satisfied

$$\text{Example: } \begin{bmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{bmatrix}$$

- If A is an orthonormal matrix then:

$$AA^T = A^T A = I \quad \text{(i.e., } A^{-1} = A^T)$$

$$||Av|| = ||v|| \text{ (does not change the magnitude of } v)$$