Linear Algebra Review

Appendix A.2 (Duda et al.)

CS479/679 Pattern Recognition
Dr. George Bebis

n-dimensional vector

- An *n*-dimensional vector *v* *is* represented as an ordered list of values in column format:

$$v = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

- The transpose v^T is denoted in row format:

$$v^T = [\, x_1 \; x_2 \; \cdots \; x_n \,]$$

Dot product

- Given $v^T = (x_1, x_2, \ldots, x_n)$ and $w^T = (y_1, y_2, \ldots, y_n)$, their dot product defined as follows:

$$v.w = x_1y_1 + x_2y_2 + \ldots + x_ny_n$$ (scalar)

or $$v.w = [\, x_1 \; x_2 \; \cdots \; x_n \,] \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix} = v^T w$$

Orthogonal / Orthonormal vectors

- A set of vectors $x_1, x_2, . . . , x_n$ is *orthogonal* if

$$x_i^T x_j = \begin{cases} k & if\ i = j \\ 0 & if\ i \neq j \end{cases}$$

length of x_i=k

- A set of vectors $x_1, x_2, . . . , x_n$ is *orthonormal* if

$$x_i^T x_j = \begin{cases} 1 & if\ i = j \\ 0 & if\ i \neq j \end{cases}$$

length of x_i=1

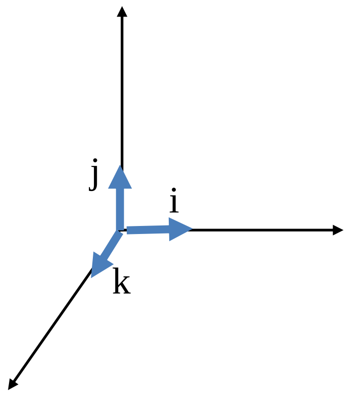
Linear combinations

- A vector v is a linear combination of the vectors $v_1, \ldots, v_n$ if:

$$v = c_1 v_1 + c_2 v_2 + \ldots + c_n v_n$$

where $c_1, \ldots, c_k$ are constants.

Example: any vector in R^3 can be expressed as a linear combination of the unit vectors $i = (1, 0, 0)$, $j = (0, 1, 0)$, and $k = (0, 0, 1)$



$$v = (a, b, c) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1)$$

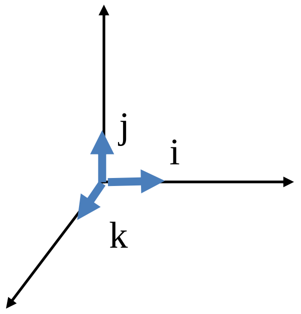
Space spanning

- A set of vectors $S=(v_1, v_2, \ldots, v_n)$ *span* some space W if every vector v in W can be written as a linear combination of the vectors in S

$$v = c_1 v_1 + c_2 v_2 + \ldots + c_n v_n$$

Example: the unit vectors *i*, *j*, and *k* span R^3

$$v = (a, b, c) = a(1,0,0) + b(0,1,0) + c(0,0,1)$$



Linear dependence

- A set of vectors $v_1, \ldots, v_n$ are *linearly dependent* if at least one of them (e.g., v_j) can be written as a linear combination of the rest:

$$v_j = c_1 v_1 + c_2 v_2 + \ldots + c_{j-1} v_{j-1} + c_{j+1} v_{j+1} + \ldots + c_n v_n$$

(i.e., v_j does not appear on the right side of the equation above)

Example: $x_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ $x_2 = \begin{bmatrix} 4 \\ 1 \\ 5 \end{bmatrix}$ $x_3 = \begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix}$ are *linearly dependent*

$$x_3 = -2x_1 + x_2.$$

Linear independence

- A set of vectors $v_1, \ldots, v_n$ is *linearly independent* if <u>no</u> vector v_j from $v_1, \ldots, v_n$ can be represented as a linear combination of the remaining vectors, i.e. :

$$c_1 v_1 + c_2 v_2 + \ldots + c_n v_n = 0 \Rightarrow c_1 = c_2 = \ldots = c_n = 0$$

<u>Example</u>: $v_1 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ are *linearly independent*

Let $c_1 v_1 + c_2 v_2 = 0$, then $\begin{bmatrix} -c_1 + c_2 \\ c_1 + c_2 \\ -c_1 + (-c_2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

This can only be true if $c_1=c_2=0$

Vector basis

- A set of vectors $v_1, \ldots, v_n$ forms a *basis* in some vector space W if:
  - (1) $(v_1, \ldots, v_n)$ span W
  - (2) $(v_1, \ldots, v_n)$ are linearly independent

Some standard bases:

R^2	R^3	R^n
$i = (1, 0),\ j = (0, 1)$	$i = (1, 0, 0),\ j = (0, 1, 0),\ k = (0, 0, 1)$	$(1,0,\ldots,0)\ (0,1,\ldots,0),\ \ldots.,\ (0,0,\ldots,1)$

Orthonormal vector basis

- Basis vectors are both orthogonal and unit length.
- Very useful in practice since they **simplify** calculations.

- Any set of basis vectors (v_1, ..., v_n) can be transformed to an orthogonal basis using the *Gram-Schmidt* orthogonalization algorithm.

- Normalizing the basis vectors to "unit" length will yield an orthonormal basis.

Vector Expansion (or Projection)

- Suppose $v_1, v_2, \ldots, v_n$ is an orthogonal base in W, then any $v \in W$ can be represented in this basis as follows:

$$v = \sum_{i=1}^{n} x_i v_i = x_1 v_1 + x_2 v_2 + \ldots + x_n v_n$$

(x_i: coefficients of expansion or projection coefficients)

- In essence, every basis $v_1, v_2, \ldots, v_n$ defines a coordinate system in W.

What is the representation of v in this coordinate system?

$$v = \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix}$$

Vector Expansion (or Projection) (cont'd)

- How do we find the expansion or projection coefficients x_i?

$$v = \sum_{i=1}^{n} x_i v_i = x_1 v_1 + x_2 v_2 + \ldots + x_n v_n$$

$$(v.v_i) = x_1(v_1.\ v_i) + x_2(v_2.v_i) + \ldots + x_i(v_i.v_i) + \ldots + x_n(v_n.v_i) = x_i(v_i.v_i)$$

$$(v.v_i) = x_i(v_i.v_i) \quad or \quad x_i = \frac{(v.v_i)}{(v_i.v_i)}$$

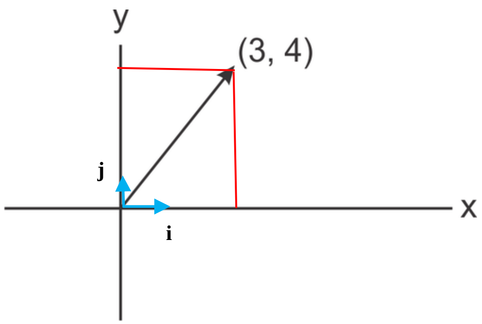
- If the base is orthonormal, then $(v_i.v_i)=1$

$$x_i = (v.v_i)$$

Example

- Let's assume n=2:

$$\mathbf{x}:\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 3 \\ 4 \end{bmatrix}$$



- Assuming the standard base <v_1=i, v_2=j>, the projection coefficients x_i can be obtained by projecting **x** along the direction of v_1 and v_2 :

$$x_1=\mathbf{x}^T i=\begin{bmatrix} 3 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}=3$$

$$x_2=\mathbf{x}^T j=\begin{bmatrix} 3 & 4 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}=4$$

- The expansion of x onto (v_1,v_2) is:

$$\mathbf{x}=3i+4j$$

Vector Expansion (or Projection) (cont'd)

- Given a set of basis vectors in W, each vector in W can be represented (i.e., projected) "uniquely" in this basis.
- How many vector bases are there?
  - Many, for example, translate/rotate a given vector basis to obtain a new one!
  - Some vector bases are preferred over others.
  - We will revisit this issue when we discuss **Principal Component Analysis** (PCA).

Matrix Operations

- Matrix addition/subtraction
  - Add/Subtract corresponding elements.
  - Matrices must be of the same size.

- Matrix multiplication

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

Condition: **n = q**

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

Not commutative: $AB \neq BA$

Diagonal Matrices

$$A = \begin{bmatrix} a_{11} & 0 & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ 0 & 0 & a_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix}$$

Special case: identity matrix

$$I = \begin{bmatrix} 1 & 0 & . & 0 \\ 0 & 1 & . & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & . & 1 \end{bmatrix} \qquad AI = IA = A,$$

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 Matrices

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

Example:

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

Determinants

2 x 2

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

3 x 3

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

(expanded along 1st column)

n x n

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

(expanded along kth column)

Properties:

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

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

Matrix Inverse

- The inverse A^{-1} of a matrix A, satisfies the property:

$$A\,A^{-1} = A^{-1}A = I$$

- A^{-1} exists only if:

$$det(A) \neq 0$$

- If A^{-1} does not exist, then we call A **singular** matrix.
- If A^{-1} exists, then we call A **non-singular** matrix.

Matrix Inverse (cont'd)

- Useful properties:

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

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

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

- A is orthogonal if:

$$AA^T = A^TA = I \quad \underbrace{\begin{bmatrix} a_{11} & a_{12} & . & a_{1n} \\ a_{21} & a_{22} & . & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & . & a_{mn} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a_{11} & a_{21} & . & a_{m1} \\ a_{12} & a_{22} & . & a_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & . & a_{mn} \end{bmatrix}}_{A^T} = \underbrace{\begin{bmatrix} 1 & 0 & . & 0 \\ 0 & 1 & . & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & . & 1 \end{bmatrix}}_{I}$$

- What is the inverse of an orthogonal matrix?

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

Matrix Rank

- The rank of a matrix A is defined as the **size** of the largest square sub-matrix of *A* whose determinant is non-zero.

Example:
$$\begin{bmatrix} 4 & 5 & 2 & 14 \\ 3 & 9 & 6 & 21 \\ 8 & 10 & 7 & 28 \\ 1 & 2 & 9 & 5 \end{bmatrix}$$ has rank 3!

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

Matrix Rank (cont'd)

- **Alternatively,** it can be defined as the maximum number of linearly independent columns (or rows) of A.

Example: $\begin{bmatrix} 4 & 5 & 2 & 14 \\ 3 & 9 & 6 & 21 \\ 8 & 10 & 7 & 28 \\ 1 & 2 & 9 & 5 \end{bmatrix}$ has rank 3!

Its columns are **not** linearly independent!

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

Matrix Rank (cont'd)

- Useful properties:

If A is nxn, $rank(A) = n$ iff A is nonsingular

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

full rank

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

rank-deficient

Eigenvalues and Eigenvectors

- The vector **v** is an eigenvector of an nxn matrix *A* and **λ** is an eigenvalue of *A* if:

$$Av = \lambda v$$ (assuming v is non-zero)

Geometric interpretation: the linear transformation implied by *A* cannot change the direction of the eigenvectors v, only their magnitude.

Computing λ and v

- To compute the eigenvalues λ of a matrix *A*, find the roots of the characteristic polynomial.

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

Example:

$$A = \begin{bmatrix} 5 & -2 \\ 6 & -2 \end{bmatrix} \Rightarrow det(\begin{bmatrix} 5-\lambda & -2 \\ 6 & -2-\lambda \end{bmatrix}) = 0 \text{ or } \lambda^2 - 3\lambda + 2 = 0 \text{ or } \lambda_1 = 1,\ \lambda_2 = 2$$

- Then, the eigenvectors can be computed:

$$Av = \lambda v \qquad v_1 = \begin{bmatrix} 1/2 \\ 1 \end{bmatrix},\ v_2 = \begin{bmatrix} 2/3 \\ 1 \end{bmatrix}$$

Properties of λ and v

- If A is **symmetric**, its eigenvalues are real and non-negative.
- Eigenvectors are not unique (e.g., if *v* is an eigenvector, so is *kv*) → $Av = \lambda v \ or \ A(kv) = \lambda(kv)$
- If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of *A*, then:

$$\prod_i \lambda_i = det(A)$$

- If $\lambda_i=0$, then A is singular!

Matrix diagonalization

- An nxn matrix A is diagonalizable, if there exist matrices P and Λ such that:

  $P^{-1}AP=\Lambda$ where Λ *is* **diagonal**

- Set $P = [v_1\ v_2\ .\ .\ .\ v_n]$, where $v_1, v_2, .\ .\ .\ v_n$ are the eigenvectors of A:

eigenvalues of A

$$
\begin{aligned}
Av_1 &= \lambda_1 v_1 \\
Av_2 &= \lambda_2 v_2 \\
&\ \vdots \\
Av_n &= \lambda_n v_n
\end{aligned}
\quad \Rightarrow \quad
AP = P\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 0 & 0 & \lambda_n \end{bmatrix}
\quad \Rightarrow \quad
P^{-1}AP = \underbrace{\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 0 & 0 & \lambda_n \end{bmatrix}}_{\Lambda}
$$

Matrix diagonalization (cont'd)

Example: $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

$$\lambda_1 = 0,\ \lambda_2 = 2,\ v_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix},\ v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

$$P = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix},\ P^{-1} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

$$P^{-1}AP = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

Are all n x n matrices diagonalizable?

$$P^{-1}AP=\Lambda$$

- An *nxn* matrix *A* is diagonalizable **iff** A has *n* linearly independent eigenvectors or rank(P)=n.
- If the eigenvalues of *A* are all distinct, then the corresponding eigenvectors are linearly independent (i.e., *A* is diagonalizable).
- **Symmetric** matrices are always diagonalizable.

https://www.statlect.com/matrix-algebra/matrix-diagonalization
https://www.statlect.com/matrix-algebra/algebraic-and-geometric-multiplicity-of-eigenvalues
(formal definition based on **algebraic** and **geometric** multiplicity of eigenvalues)

Very Important Property!!!

- If A is diagonalizable, then its eigenvectors $v_1, v_2, \ldots v_n$ form a basis in R^n

- If A is symmetric, its eigenvectors are also orthogonal.

$v_1, v_2, \ldots v_n$ form an orthogonal basis in R^n

Matrix eigen-decomposition

- If *A* is diagonalizable, then A can be decomposed as follows:

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ . & . & . \\ . & . & . \\ . & . & . \\ 0 & 0 & \lambda_n \end{bmatrix} \quad \Rightarrow \quad A = P \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ . & . & . \\ . & . & . \\ . & . & . \\ 0 & 0 & \lambda_n \end{bmatrix} P^{-1}$$

Matrix eigen-decomposition (cont'd)

- If A is symmetric, matrix decomposition can be further simplified:

$$A = P \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ . & . & . \\ . & . & . \\ . & . & . \\ 0 & 0 & \lambda_n \end{bmatrix} P^{-1}$$

P is orthogonal:
$P^{-1}=P^T$

$\Rightarrow$ $A=PDP^T= \sum_{i=1}^{n} \lambda_i v_i v_i^T$