



College of Engineering & Applied Sciences

# CSPB 2820

*Linear Algebra With Computer Science Applications*

*Exam Notes*

TAYLOR LARRECHEA

2023

## Exam 1 Notes

**Inner Product:** Calculates the overlap between two vectors

$$a^T b = \|a\| \|b\| \cos(\theta).$$

When  $\theta = \pi/2$  the vectors are perpendicular and when  $\theta = 0$  the vectors are parallel.

**Taylor Series:** Approximates a differentiable function to a polynomial at a given point of the original function

$$\begin{aligned} \mathbb{T}(x, a) &= \sum_{n=0}^{\infty} \frac{f^n(x-a)^n}{n!} \\ \hat{f}(x) &= f(x, z) + \nabla f(z)^T (x - z) \\ \nabla f(x, z) &= \begin{bmatrix} \frac{\partial f}{\partial x_1}(z) \\ \vdots \\ \frac{\partial f}{\partial x_n}(z) \end{bmatrix}. \end{aligned}$$

**Affine Function:** Inner product with an additional scalar value

$$f(x) = a^T x + b.$$

**Linear Functions:** Linear functions are of the form

$$\text{Superposition: } f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

$$\text{Homogeneity: } f(\alpha x) = \alpha f(x)$$

$$\text{Additivity: } f(x + y) = f(x) + f(y).$$

Superposition is a combination of homogeneity and additivity.

**Linear Regression:** Models the relationship between a dependent variable and one or more independent variables, aiming to estimate coefficients that minimize the error between predicted and observed values

$$\hat{y} = x^T \beta + v$$

where  $x$  are the regressors,  $\hat{y}$  is the prediction,  $\beta$  is an  $n$ -vector and  $v$  is a scalar.

**Norm Of Vector:** Measurement of how large a vector or a collection of vectors are

$$\begin{aligned} \|x\| &= x^T x = \sqrt{x_1^2 + \dots + x_n^2} \\ \|x + y\| &= \sqrt{\|x\|^2 + 2x^T y + \|y\|^2} \\ \|(a, b, c)\| &= \sqrt{\|a\|^2 + \|b\|^2 + \|c\|^2}. \end{aligned}$$

**Triangle Inequality:** The norm of a sum of two vectors is no more than the sum of their norms

$$\|x + y\| \leq \|x\| + \|y\|.$$

**Root Mean Square (RMS):** Typical value of  $\|x_i\|$

$$\text{rms}(x) = \frac{\|x\|}{\sqrt{n}}.$$

**Chebyshev Inequality:** Provides an upper bound on the probability that a random variable deviates from its mean by a certain multiple of its standard deviation

$$\frac{k}{n} \leq \left( \frac{\text{rms}(x)}{a} \right)^2$$

where  $k$  is the number of entries in  $x$  with absolute value of at least  $a$ .

**Distance:** Calculates the Euclidean distance between two vectors

$$\text{dist}(a, b) = \|a - b\|.$$

**Average:** Calculates the average value of a vector

$$\text{avg}(x) = \frac{x_1 + \dots + x_n}{n}.$$

**Nearest Neighbor:** The vector that is the closest vector to another given vector

$$\|x - z_j\| \leq \|x - z_i\|, \quad i = 1, \dots, n.$$

**Standard Deviation:** A measure of the dispersion of the values within the vector

$$\text{std}(x) = \frac{\|x - (\mathbf{1}^T x / n)\|}{\sqrt{n}}.$$

**J-Clust:**  $J_C$  is a method of quantifying how good a clustering of data points is for a given cluster

$$J_C = \sum_{i=1}^N \sum_{j=1}^k \frac{1}{N} \min(\|x_i - z_j\|^2)$$

where  $z_j$  is calculated by

$$z_j = \frac{1}{|G_j|} \sum_{i \in G_j} x_i.$$

**K-Means Algorithm:** The  $k$ -means algorithm clusters data for a given data set into  $k$  clusters.

1. *Partition the vectors into  $k$  groups:* For each vector  $i = 1, \dots, N$ , assign  $x_i$  to the group associated with the nearest representative
2. *Update representatives:* For each group  $j = 1, \dots, k$ , set  $z_j$  to be the mean of the vectors in group  $j$ .

**Linear Independence:** Linear independence and dependence can be summarized with

$$\beta_1 a_1 + \dots + \beta_k a_k = 0$$

where for a collection of vectors to be linearly independent or dependent one of the following statements must be true.

1. *Linearly Independent:* All values of  $\beta$  must be zero.
2. *Linearly Dependent:* There exists at least one  $\beta$  that is non zero.

## Exam 2 Notes

**Matrix Types:** Matrix sizes are defined in terms of their rows ( $m$ ) by columns ( $n$ ):

- **Square Matrix:**  $m = n$ .
- **Tall Matrix:**  $m > n$ .
- **Wide Matrix:**  $m < n$ .

**2D Rotation Matrix:** The 2D rotation matrix is:

$$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

**3D Rotation Matrices:** The 3D rotation matrices are:

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix}$$

$$R_y(\theta) = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix}$$

$$R_z(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

**Incidence Matrix:** An incidence matrix is of the form:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

The rows and columns of an incidence matrix represent:

- **Columns:** Represent the edges of the graph.
- **Rows:** Represent the vertices of the graph.

The values of the elements indicate:

- 1: Edge  $j$  points to node  $i$ .
- -1: Edge  $j$  points from node  $i$ .
- 0: Otherwise.

**Dirichlet Energy:** Measures the potential differences across edges:

$$\mathcal{D}(v) = \|A^T\|^2 = \sum_{(k,l) \in \text{Edges}} (v_l - v_k)^2$$

**Matrix Multiplication:** Matrix multiplication is performed by:

$$\begin{aligned} AB &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} \end{aligned}$$

The main properties of matrix multiplication are:

- **Associativity:**  $(AB)C = A(BC)$
- **Scalar Multiplication:**  $\gamma(AB) = (\gamma A)B$
- **Distributivity:**

$$- A(B + C) = AB + AC$$

$$- (A + B)C = AC + BC$$

- **Transpose Product:**  $(AB)^T = B^T A^T$

**Some Matrix Representations:** Some matrix representations are:

- **Column Interpretation:**

$$AB = \begin{bmatrix} Ab_1 & Ab_2 & \dots & Ab_n \end{bmatrix}$$

- **Row Interpretation:**

$$AB = \begin{bmatrix} (B^T a_1)^T \\ (B^T a_2)^T \\ \vdots \\ (B^T a_m)^T \end{bmatrix}$$

- **Gram Matrix:**

$$G = A^T A$$

- **QR Factorization:**

$$A = QR$$

**Matrix Inverses:** The rules for if a matrix is invertible:

- Matrix must be square.
- The columns are linearly independent.
- The rows are linearly independent.
- The matrix has a left inverse.
- The matrix has a right inverse.

For a  $2 \times 2$  matrix the inverse is:

$$A^{-1} = \frac{1}{|A|} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

For a  $3 \times 3$  matrix the inverse is:

$$A^{-1} = \frac{1}{|A|} \begin{bmatrix} \left\| \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \right\| & \left\| \begin{bmatrix} a_{13} & a_{12} \\ a_{33} & a_{32} \end{bmatrix} \right\| & \left\| \begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{bmatrix} \right\| \\ \left\| \begin{bmatrix} a_{23} & a_{21} \\ a_{33} & a_{31} \end{bmatrix} \right\| & \left\| \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} \right\| & \left\| \begin{bmatrix} a_{13} & a_{11} \\ a_{23} & a_{21} \end{bmatrix} \right\| \\ \left\| \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \right\| & \left\| \begin{bmatrix} a_{12} & a_{11} \\ a_{32} & a_{31} \end{bmatrix} \right\| & \left\| \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \right\| \end{bmatrix}$$