2. Vector Space

Chapter 1 - 3 of Rancher & Schaalje

Euclidean space:

A vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ of order 2 represents a point in a plane

Note that any point in the plane can be represented as

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

basis vectors

The entire plane is denoted by R^2.

A vector of order 3, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ represents a point in 3-dimensional Euclidean space (denoted by R^3).

Note that any $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in R^3$ can be expressed as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

basis vectors for R^3

A vector of order n, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ represents a point in n-dimensional Euclidean space (denoted by R^n).

R^n is a special case of a more general concept of a **vector space.**

Defn 2.1: A set of vectors, denoted by S, is a **vector space** if for every pair of vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ in S we have

(i) $\mathbf{x}_i + \mathbf{x}_j$ is a vector in S
(ii) $a\mathbf{x}_i$ is in S for any real scalar.

Defn 2.2: If every vector in some vector space S can be expressed as a linear combination

$$a_1\,\mathbf{x}_1 + a_2\,\mathbf{x}_2 + \cdots + a_k\,\mathbf{x}_k$$

of a set of k vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$, this set of vectors is said to **span** the vector space S.

Defn 2.3: If a set of vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$ span S and are linearly independent, then the set is called a **basis** for S.

Comments:

(i) The number of vectors in a basis for a vector space S is called the dimension of S ($\dim(S)$).

(ii) $\mathbf{0}$ belongs to every vector space in R^n.

(iii) A vector space can have many bases.

Example:

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \mathbf{x}_4 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

span R^3, but are <u>not</u> a basis for R^3.

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

are a basis for R^3.

Note that

$$\frac{1}{3}\mathbf{x}_1 + \frac{1}{3}\mathbf{x}_2 + \frac{1}{3}\mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\frac{1}{3}\mathbf{x}_1 - \frac{2}{3}\mathbf{x}_2 + \frac{1}{3}\mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\frac{1}{3}\mathbf{x}_1 + \frac{1}{3}\mathbf{x}_2 - \frac{2}{3}\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

then

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$= \left(\frac{a+b+c}{3} \right) \mathbf{x}_1$$

$$+ \left(\frac{a-2b+c}{3} \right) \mathbf{x}_2$$

$$+ \left(\frac{a+b-2c}{3} \right) \mathbf{x}_3$$

Example:

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \; \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \; \mathbf{x}_3 = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

do not span R^3. Any two of these vectors provides a basis for a 2-dimensional subspace of R^3.

Note that $\mathbf{x}_3 = \mathbf{x}_1 + 2\,\mathbf{x}_2$, which implies that $\mathbf{x}_1 = \mathbf{x}_3 - 2\,\mathbf{x}_2$ and $\mathbf{x}_2 = 0.5\,(\mathbf{x}_3 - \mathbf{x}_1)$.

Then, for any $\mathbf{z} = a\,\mathbf{x}_1 + b\,\mathbf{x}_2$, we have

$$\begin{aligned} \mathbf{z} &= a(\mathbf{x}_3 - 2\,\mathbf{x}_2) + b\,\mathbf{x}_2 \\ &= (b - 2a)\mathbf{x}_2 + a\,\mathbf{x}_3 \end{aligned}$$

and

$$\begin{aligned}\mathbf{z} &= a\mathbf{x}_1 + \frac{b}{2}(\mathbf{x}_3 - \mathbf{x}_1) \\ &= (a - \frac{b}{2})\mathbf{x}_1 + \frac{b}{2}\mathbf{x}_3\end{aligned}$$

This 2-dimensional subspace of R^3 is the vector space consisting of all vectors of the form

$$\begin{aligned}\mathbf{z} = a\mathbf{x}_1 + b\mathbf{x}_2 &= a\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + b\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} a+b \\ 2a \\ b-a \end{bmatrix}\end{aligned}$$

Random vectors:

Defn 2.4: A random vector $\mathbf{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}$ is a vector whose elements are random variables.

Mean vectors:

$$E(\mathbf{Y}) = \begin{bmatrix} E(Y_1) \\ \vdots \\ E(Y_n) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = \boldsymbol{\mu}$$

where

$$\mu_i = E(Y_i) = \begin{cases} \int_{-\infty}^{\infty} y\, f_i(y) dy & \text{if } Y_i \text{ is a continuous random variable with density function } f_i(y) \\ \displaystyle\sum_{\substack{all\ \ possible \\ y\ \ values}} y\, p_i(y) & \text{if } Y_i \text{ is a discrete random variable with probability function } p_i(y). \end{cases}$$

Covariance matrix:

$$\Sigma = Var(\mathbf{Y}) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \cdots & \sigma_n^2 \end{bmatrix}$$

with variances

$$Var(Y_i) = \sigma_i^2 = E(Y_i - \mu_i)^2$$

$$
= \begin{cases} \int_{-\infty}^{\infty} (y - \mu_i)^2 f_i(y) dy & \begin{array}{l} \text{if } y_i \text{ is a continuous} \\ \text{random variable} \end{array} \\ \sum_{\substack{all \\ y}} (y - \mu_i)^2 \, p_i(y) & \begin{array}{l} \text{if } y_i \text{ is a discrete} \\ \text{random variable} \end{array} \end{cases}
$$

and covariances:

$$\sigma_{ij} = Cov(Y_i, Y_j) = E\left[(Y_i - \mu_i)(Y_j - \mu_j)\right]$$

where $$\sigma_{ij} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(y-\mu_i)(v-\mu_j)f_{ij}(y,v)dy\,dv$$

if Y_i and Y_j are continuous random variables with joint density function $f_{ij}(y, v)$

and $$\sigma_{ij} = \sum_{\substack{\text{all}\\ y}}\sum_{\substack{\text{all}\\ v}}(y-\mu_i)(v-\mu_j)P_{ij}(y,v)$$

if Y_i and Y_j are discrete random variables with joint probability function $P_{ij}(y, v) = Pr(Y_i = y, V_j = v)$

Result 2.1:

Let $\mathbf{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}$ be a random vector with

$$\mu = E(\mathbf{Y}) \text{ and } \sum = Var(\mathbf{Y}),$$

and let

$$A_{p\times n} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{p1} & \cdots & a_{pn} \end{bmatrix}$$

be a matrix of non-random elements,

and let

$$\mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \text{ and } \mathbf{d} = \begin{bmatrix} d_1 \\ \vdots \\ d_p \end{bmatrix}$$

be vectors of non-random elements, then

(i) $E(A\mathbf{Y} + \mathbf{d}) = A\mu + \mathbf{d}$

(ii) $Var(A\mathbf{Y} + \mathbf{d}) = A\Sigma A^T$

(iii) $E(\mathbf{c}^T\mathbf{Y}) = \mathbf{c}^T\mu$

(iv) $Var(\mathbf{c}^T\mathbf{Y}) = \mathbf{c}^T\Sigma\mathbf{c}$