II. Foundations of Multivariate Analysis

A Some Matrix Algebra

Partitioned Matrices

$$\underset{(m+q)\times(n+r)}{\mathbf{A}} = \begin{bmatrix} \underset{m\times n}{\mathbf{A}_{11}} & \underset{m\times r}{\mathbf{A}_{12}} \\ \underset{q\times n}{\mathbf{A}_{21}} & \underset{q\times r}{\mathbf{A}_{22}} \end{bmatrix} \quad \underset{(n+r)\times(s+t)}{\mathbf{B}} = \begin{bmatrix} \underset{n\times s}{\mathbf{B}_{11}} & \underset{n\times t}{\mathbf{B}_{12}} \\ \\ \underset{r\times s}{\mathbf{B}_{21}} & \underset{r\times t}{\mathbf{B}_{22}} \end{bmatrix}.$$

– Since all of the submatrices are conformable,

$$\mathbf{AB} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{bmatrix}$$

– If $\mathbf{A}_{11}$ and $\mathbf{A}_{22}$ are square ($m = n$ and $q = r$) and $\mathbf{A}_{12} = \mathbf{A}'_{21} = \mathbf{0}$,

$$|\mathbf{A}| = |\mathbf{A}_{11}|\,|\mathbf{A}_{22}|$$

– If $\mathbf{A}_{11}$ and $\mathbf{A}_{22}$ are both square and nonsingular,

$$\begin{aligned} |\mathbf{A}| &= |\mathbf{A}_{11}| \left| \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \right| \\ &= |\mathbf{A}_{22}| \left| \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} \right| \end{aligned}$$

analogous to

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} &= a_{11}a_{22} - a_{21}a_{12} \\ &= a_{11}\left(a_{22} - \frac{a_{21}a_{12}}{a_{11}}\right) \\ &= a_{22}\left(a_{11} - \frac{a_{12}a_{21}}{a_{22}}\right) \end{aligned}$$

Eigenvalues and Eigenvectors

λ is an eigenvalue of the square matrix $\mathbf{A}$ and $\mathbf{x}$ is the corresponding eigenvector if:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

or

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$$

- If $|\mathbf{A} - \lambda\mathbf{I}| \neq 0$, then $(\mathbf{A} - \lambda\mathbf{I})$ has an inverse and

$$(\mathbf{A} - \lambda\mathbf{I})^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = (\mathbf{A} - \lambda\mathbf{I})^{-1}\ \mathbf{0}$$
$$\Rightarrow \mathbf{x} = \mathbf{0} \text{ is the only solution}$$

- So, set $\underbrace{|\mathbf{A} - \lambda\mathbf{I}_p| = 0}_{\text{“characteristic equation”}}$ and solve for λ

- Eigenvalues $\lambda_1, \ldots, \lambda_p$ accompanied by eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_p$.

Orientation of eigenvectors is what's important. Length is arbitrary — $k\mathbf{x}_1, \ldots, k\mathbf{x}_p$ equally good. (We usually choose eigenvectors such that $\mathbf{x}'\mathbf{x} = 1$.)

- Spectral decomposition of $\mathbf{A}$

  Let

$$\begin{aligned} \mathbf{C} &= \text{matrix containing normalized eigenvectors of } \mathbf{A} \\ &= [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \ldots \quad \mathbf{x}_p] \end{aligned}$$

  and let

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{bmatrix}$$

  Note that $\mathbf{C}$ is orthogonal (so $\mathbf{I} = \mathbf{C}\mathbf{C}' = \mathbf{C}'\mathbf{C}$)

$$
\begin{aligned}
\mathbf{A} &= \mathbf{ACC}' \\
&= \mathbf{A}[\mathbf{x}_1 \quad \mathbf{x}_2 \quad \ldots \quad \mathbf{x}_p]\mathbf{C}' \\
&= [\mathbf{Ax}_1 \quad \mathbf{Ax}_2 \quad \ldots \quad \mathbf{Ax}_p]\mathbf{C}' \\
&= [\lambda_1\mathbf{x}_1 \quad \lambda_2\mathbf{x}_2 \quad \ldots \quad \lambda_p\mathbf{x}_p]\mathbf{C}' \\
&= \mathbf{CDC}' \\
&= \lambda_1\mathbf{x}_1\mathbf{x}_1' + \lambda_2\mathbf{x}_2\mathbf{x}_2' + \cdots + \lambda_p\mathbf{x}_p\mathbf{x}_p'
\end{aligned}
$$

Also,

$$\mathbf{D} = \mathbf{C}'\mathbf{AC}$$

Positive Definite & Nonnegative Definite Matrices

- $\underset{p \times p}{\mathbf{A}}$ is symmetric
- If $\mathbf{x}'\mathbf{A}\mathbf{x} > \mathbf{0}$ for all $\mathbf{x} \neq \mathbf{0}$, then $\mathbf{A}$ is "positive definite (p.d.)"
- If $\mathbf{x}'\mathbf{A}\mathbf{x} \geq \mathbf{0}$ for all $\mathbf{x} \neq \mathbf{0}$, then $\mathbf{A}$ is "nonnegative definite (n.n.d.)"
- If $\mathbf{x}'\mathbf{A}\mathbf{x} \geq \mathbf{0}$ for all $\mathbf{x} \neq \mathbf{0}$, with $\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{0}$ for at least one $\mathbf{x} \neq \mathbf{0}$, then $\mathbf{A}$ is "positive semi-definite (p.s.d.)"
  - In other words, if $\mathbf{A}$ is n.n.d., but not p.d., we say $\mathbf{A}$ is p.s.d.

Easy check for these properties:

1. Eigenvalues of a positive definite matrix are all positive.
2. Eigenvalues of a nonnegative definite matrix are positive or zero (with rank($\mathbf{A}$) = number of positive eigenvalues).

Trace and Determinant of a Square Matrix $\underset{p\times p}{\mathbf{A}} = (a_{ij})$

- $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{p} a_{ii}$
- $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$

If $\mathbf{A}$ has e'vals $\lambda_1, \ldots, \lambda_p$

- $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{p} \lambda_i$
- $|\mathbf{A}| = \prod_{i=1}^{p} \lambda_i$ (Practice: show these two statements are true)

Square-root and Inverse matrices

The spectral decomposition of symmetric $\underset{p\times p}{\mathbf{A}}$:

$$\mathbf{A} = \mathbf{C}\mathbf{D}\mathbf{C}'$$

- $\mathbf{A}^{1/2} = \mathbf{C}\mathbf{D}^{1/2}\mathbf{C}'$

$$\text{where } \mathbf{D}^{1/2} = \begin{bmatrix} \lambda_1^{1/2} & & & \mathbf{0} \\ & \lambda_2^{1/2} & & \\ & & \ddots & \\ \mathbf{0} & & & \lambda_p^{1/2} \end{bmatrix}$$

$$\begin{aligned}\text{Note: } \mathbf{A}^{1/2}\mathbf{A}^{1/2} &= \mathbf{C}\mathbf{D}^{1/2}\mathbf{C}'\mathbf{C}\mathbf{D}^{1/2}\mathbf{C}' \\ &= \mathbf{C}\mathbf{D}^{1/2}\mathbf{D}^{1/2}\mathbf{C}' \\ &= \mathbf{C}\mathbf{D}\mathbf{C}' \\ &= \mathbf{A}\end{aligned}$$

- $\mathbf{A}^{-1} = \mathbf{C}\mathbf{D}^{-1}\mathbf{C}'$

$$\text{where } \mathbf{D}^{-1} = \begin{bmatrix} \frac{1}{\lambda_1} & & & \mathbf{0} \\ & \frac{1}{\lambda_2} & & \\ & & \ddots & \\ \mathbf{0} & & & \frac{1}{\lambda_p} \end{bmatrix}$$

$$
\begin{aligned}
\text{Note: } \mathbf{A}^{-1}\mathbf{A} &= \mathbf{C}\mathbf{D}^{-1}\mathbf{C}'\mathbf{C}\mathbf{D}\mathbf{C}' \\
&= \mathbf{C}\mathbf{D}^{-1}\mathbf{D}\mathbf{C}' \\
&= \mathbf{C}\mathbf{C}' \\
&= \mathbf{I}
\end{aligned}
$$

- Suppose $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{a}_{12} \\ \mathbf{a}'_{12} & a_{22} \end{bmatrix}$

$$\mathbf{A}^{-1} = \frac{1}{b} \begin{bmatrix} b\mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{a}_{12}\mathbf{a}'_{12}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{a}_{12} \\ -\mathbf{a}'_{12}\mathbf{A}_{11}^{-1} & 1 \end{bmatrix}$$

  where $b = a_{22} - \mathbf{a}'_{12}\mathbf{A}_{11}^{-1}\mathbf{a}_{12}$

- Suppose $\underset{p\times p}{\mathbf{A}} = \mathbf{B} + \mathbf{c}\mathbf{c}'$

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} - \frac{\mathbf{B}^{-1}\mathbf{c}\mathbf{c}'\mathbf{B}^{-1}}{1 + \mathbf{c}'\mathbf{B}^{-1}\mathbf{c}}$$

Random Vectors and Matrices

- $\mathbf{X} = (x_{ij})$ is a random matrix (matrix of r.v.'s)
- $E\{\mathbf{X}\} = \begin{bmatrix} E\{x_{11}\} & E\{x_{12}\} & \cdots & E\{x_{1p}\} \\ E\{x_{21}\} & E\{x_{22}\} & \cdots & E\{x_{2p}\} \\ \vdots & \vdots & \ddots & \vdots \\ E\{x_{n1}\} & E\{x_{n2}\} & \cdots & E\{x_{np}\} \end{bmatrix}$
- If $\mathbf{X}$ and $\mathbf{Y}$ are random and $\mathbf{A}$ and $\mathbf{B}$ are constant:

$$E\{\mathbf{X} + \mathbf{Y}\} = E\{\mathbf{X}\} + E\{\mathbf{Y}\}$$

$$E\{\mathbf{AXB}\} = \mathbf{A}E\{\mathbf{X}\}\mathbf{B}$$

- If $\underset{p\times 1}{\mathbf{x}}$ is a random vector

$$E\{\mathbf{x}\} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} = \underset{p\times 1}{\boldsymbol{\mu}}$$

$$\boldsymbol{\Sigma} = \mathrm{var}\{\mathbf{x}\} \qquad (\leftarrow \text{ or "cov}\{\mathbf{x}\}\text{"})$$

$$= E\{(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})'\}$$

$$= \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$$

Notes:

1. $\sigma_{ij} = \sigma_{ji}$
2. x_i indep. of x_j $\Rightarrow$ $\mathrm{cov}(x_i, x_j) = 0$

$$\mathbf{P} = \mathrm{corr}\{\mathbf{x}\} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{bmatrix}$$

Note:

$\rho_{ij} = \rho_{ji}$

- Let

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_q \\ x_{q+1} \\ \vdots \\ x_p \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \begin{array}{l} \leftarrow\ q \times 1 \\ \leftarrow\ (p-q) \times 1 \end{array}$$

$$E\{\mathbf{x}\} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}$$

$$\text{var}\{\mathbf{x}\} = \begin{bmatrix} \underbrace{\boldsymbol{\Sigma}_{11}}_{q \times q} & \underbrace{\boldsymbol{\Sigma}_{12}}_{q \times (p-q)} \\ \underbrace{\boldsymbol{\Sigma}_{21}}_{(p-q) \times q} & \underbrace{\boldsymbol{\Sigma}_{22}}_{(p-q) \times (p-q)} \end{bmatrix}$$

- Let $\underset{r\times p}{\mathbf{B}}$ and $\underset{m\times p}{\mathbf{C}}$ be constant matrices and let $\underset{r\times 1}{\mathbf{b}}$ be $\underset{m\times 1}{\mathbf{c}}$ constant vectors

$$
\begin{aligned}
E\{\mathbf{Bx}\} &= \mathbf{B}\boldsymbol{\mu} && \leftarrow r \times 1 \\
\mathrm{var}\{\mathbf{Bx}\} &= \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}' && \leftarrow r \times r \\
\mathrm{cov}\{\mathbf{Bx}, \mathbf{Cx}\} &= \mathbf{B}\boldsymbol{\Sigma}\mathbf{C}' && \leftarrow r \times m \\
E\{\mathbf{b}'\mathbf{x}\} &= \mathbf{b}'\boldsymbol{\mu} && \leftarrow \text{scalar} \\
\mathrm{var}\{\mathbf{b}'\mathbf{x}\} &= \mathbf{b}'\boldsymbol{\Sigma}\mathbf{b} && \leftarrow \text{scalar} \\
\mathrm{cov}\{\mathbf{b}'\mathbf{x}, \mathbf{c}'\mathbf{x}\} &= \mathbf{b}'\boldsymbol{\Sigma}\mathbf{c} && \leftarrow \text{scalar}
\end{aligned}
$$

B Expected values for $\bar{\mathbf{x}}$ and $\mathbf{S}$

Let $\mathbf{x}_i$, $i = 1, \ldots, n$, be an i.i.d. random sample with $E\{\mathbf{x}_i\} = \boldsymbol{\mu}$ and $\mathrm{var}\{\mathbf{x}_i\} = \boldsymbol{\Sigma}$.

$$
\begin{aligned}
E\{\bar{\mathbf{x}}\} &= \frac{1}{n}\left(E\{\mathbf{x}_1\} + E\{\mathbf{x}_2\} + \cdots + E\{\mathbf{x}_n\}\right) \\
&= \frac{1}{n}\left(n\boldsymbol{\mu}\right) \\
&= \boldsymbol{\mu}
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{var}\{\bar{\mathbf{x}}\} &= E\{(\bar{\mathbf{x}}-\boldsymbol{\mu})(\bar{\mathbf{x}}-\boldsymbol{\mu})'\} \\
&=
\end{aligned}
$$

•

Note:

$$\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'$$

$$= \sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})\mathbf{x}_i' - \left(\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})\right)\bar{\mathbf{x}}'$$

$$= \sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i' - \sum_{i=1}^{n}\bar{\mathbf{x}}\mathbf{x}_i'$$

$$= \sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i' - n\bar{\mathbf{x}}\bar{\mathbf{x}}'$$

$$E\{\mathbf{S}\} = E\left\{\frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'\right\}$$

$$=$$

$$E\{\mathbf{S}_n\} = E\{\frac{n-1}{n}\mathbf{S}\} = \frac{n-1}{n}\mathbf{\Sigma}$$

C Geometry of the Sample

Vectors

- Length of vector $\mathbf{x} = (x_1, \ldots, x_n)$

$$L_{\mathbf{x}} = \sqrt{x_1^2 + \cdots + x_n^2} = \sqrt{\mathbf{x}'\,\mathbf{x}}$$

- For some constant c,

$$L_{(c\mathbf{x})} = |c|\, L_{\mathbf{x}}$$

- Angle between vectors

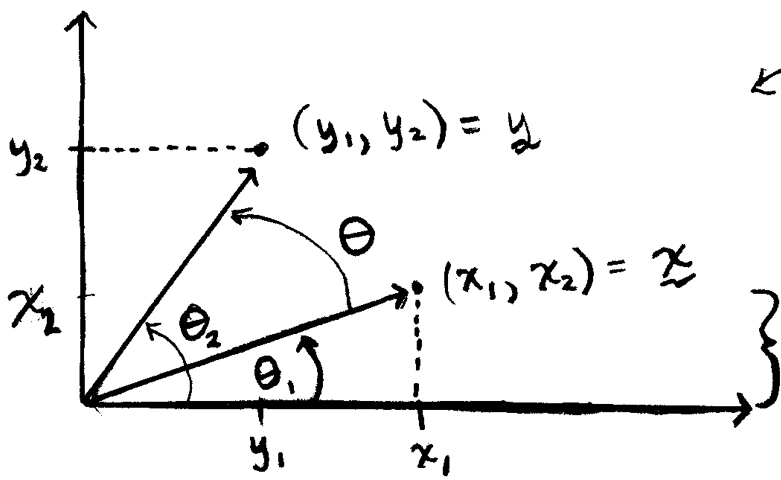


Fig. 2.4 in <u>JW</u>

$x_2 = \sin(\theta_1) \cdot L_{\underset{\sim}{x}}$

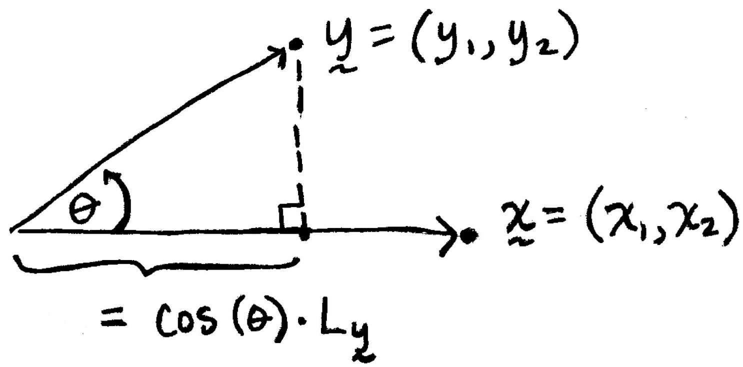
$$
\begin{aligned}
\cos(\theta) &= \cos(\theta_2 - \theta_1) \\
&= \cos(\theta_2)\cos(\theta_1) + \sin(\theta_2)\sin(\theta_1) \\
&= \left(\frac{y_1}{L_y}\right)\left(\frac{x_1}{L_{\mathbf{x}}}\right) + \left(\frac{y_2}{L_y}\right)\left(\frac{x_2}{L_{\mathbf{x}}}\right) \\
&= \frac{\mathbf{x}'\mathbf{y}}{L_{\mathbf{x}}L_{\mathbf{y}}}
\end{aligned}
$$

So

$$
\mathbf{x}'\mathbf{y} = 0 \Longleftrightarrow \cos(\theta) = 0 \Longleftrightarrow \theta \text{ is } 90^\circ \text{ or } 270^\circ
$$

Thus $\mathbf{x}'\mathbf{y} = 0$ means $\mathbf{x}$ and $\mathbf{y}$ are perpendicular.

Projections



Projection (shadow) of $\mathbf{y}$ onto $\mathbf{x}$

$$= \frac{\mathbf{y}'\mathbf{x}}{\mathbf{x}'\mathbf{x}}\mathbf{x}$$

Length of projection of $\mathbf{y}$ onto $\mathbf{x}$

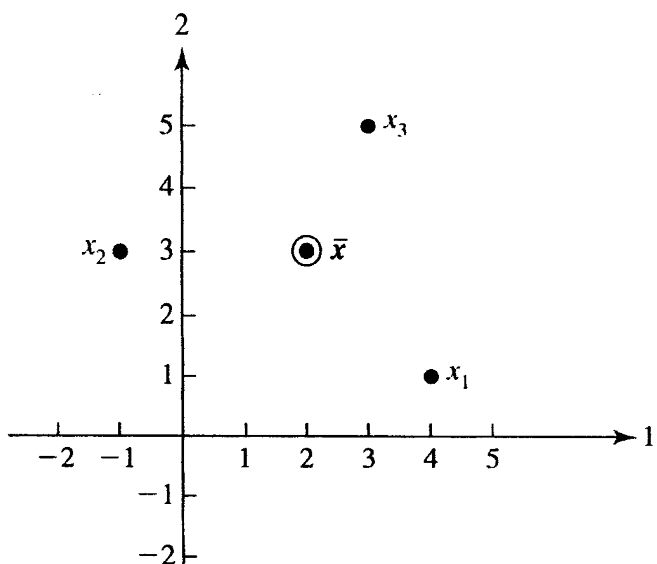
$$= \left|\frac{\mathbf{y}'\mathbf{x}}{\mathbf{x}'\mathbf{x}}\right| L_{\mathbf{x}} = \frac{|\mathbf{y}'\mathbf{x}|}{L_{\mathbf{x}}^2} L_{\mathbf{x}} = \frac{|\mathbf{y}'\mathbf{x}|}{L_{\mathbf{x}}}$$

$$= |\cos(\theta)| \cdot L_{\mathbf{y}} \qquad \text{since } \mathbf{y}'\mathbf{x} = \cos(\theta) L_{\mathbf{y}} L_{\mathbf{x}}$$

Projections of Sample Vectors

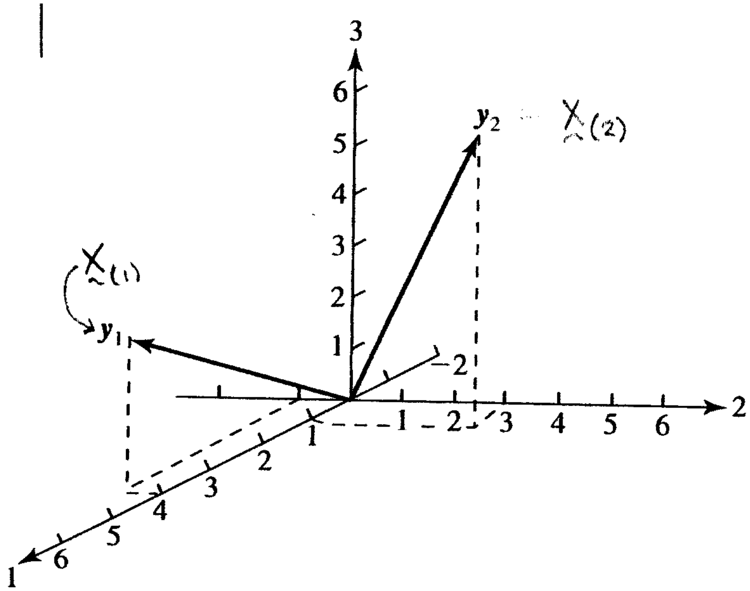
- Viewing $\underset{n\times p}{\mathbf{X}} = \begin{bmatrix} 4 & 1 \\ -1 & 3 \\ 3 & 5 \end{bmatrix}$ as n points in p-space,

$$\begin{aligned} \underset{p\times 1}{\bar{\mathbf{x}}} &= \frac{1}{n}\mathbf{1}'\mathbf{X} \\ &= (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p)' \\ &= \text{the center of gravity} \end{aligned}$$



(Fig 3.1 from JW)

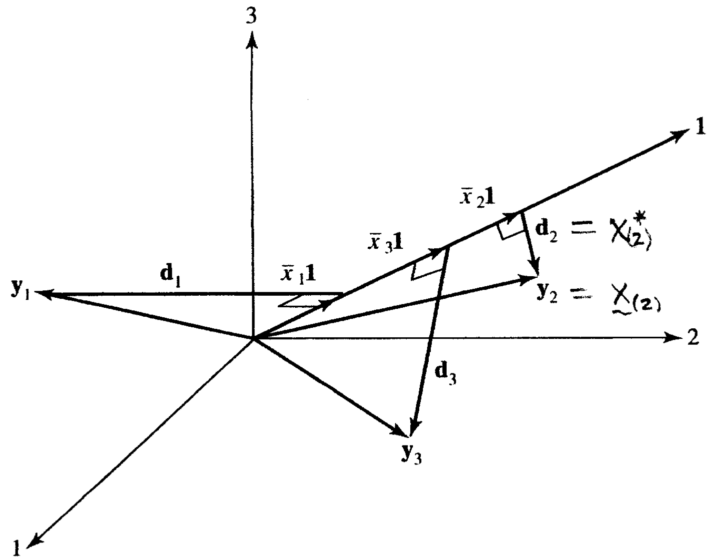
- Alternatively, view $\underset{n\times p}{\mathbf{X}} = \left[\mathbf{x}_{(1)}, \mathbf{x}_{(2)}, \cdots, \mathbf{x}_{(p)}\right]$ as p points in n-space



(Fig 3.2 from JW)

  - Projection of $\mathbf{x}_{(i)}$ onto unit-length vector $\frac{1}{\sqrt{n}}\mathbf{1}_n$ is:

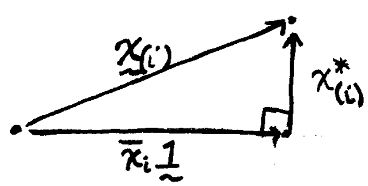
$$\frac{\mathbf{x}'_{(i)}\left(\frac{1}{\sqrt{n}}\mathbf{1}\right)}{\left(\frac{1}{\sqrt{n}}\mathbf{1}\right)'\left(\frac{1}{\sqrt{n}}\mathbf{1}\right)}\frac{1}{\sqrt{n}}\mathbf{1} = \left(\frac{1}{n}\sum_{j=1}^{n} x_{ji}\right)\mathbf{1} = \bar{x}_i\mathbf{1}$$



(Fig 3.3 from JW)

– Note: Centered (mean-corrected) version of $\mathbf{x}_{(i)}$ (also called "deviation vector") is

$$\mathbf{x}^*_{(i)} = \mathbf{x}_{(i)} - \bar{x}_i \mathbf{1}$$



– Note:

$$\begin{aligned}
\text{Length of } \mathbf{x}^*_{(i)} &= \sqrt{\mathbf{x}^{*'}_{(i)}\mathbf{x}^*_{(i)}} \\
&= \sqrt{\sum_{j=1}^{n}(x_{ji}-\bar{x}_i)^2} \\
&= \sqrt{(n-1)s_{ii}} \\
&= \sqrt{n-1}\,s_i \\
\text{OR} \qquad s_{ii} &= \frac{1}{n-1}L^2_{\mathbf{x}^*_{(i)}} \\
\text{Similarly,} \qquad s_{ij} &= \frac{1}{n-1}\mathbf{x}^{*'}_{(i)}\mathbf{x}^*_{(j)}
\end{aligned}$$

- Since, in general, $\mathbf{y}'\mathbf{z} = L_{\mathbf{y}} L_{\mathbf{z}} \cos(\theta)$, it follows that

$$
\begin{aligned}
\cos(\boldsymbol{\theta}) &= \frac{\mathbf{x}_{(i)}^{*'} \mathbf{x}_{(j)}^{*}}{L_{\mathbf{x}_{(i)}^{*}} L_{\mathbf{x}_{(j)}^{*}}} \\
&= \frac{(n-1)s_{ij}}{\sqrt{(n-1)s_{ii}}\sqrt{(n-1)s_{jj}}} \\
&= \frac{s_{ij}}{\sqrt{s_{ii}}\sqrt{s_{jj}}} \\
&= r_{ij}
\end{aligned}
$$

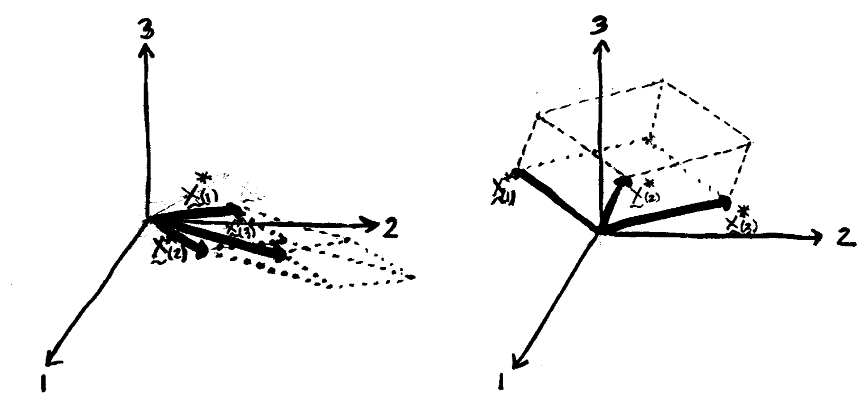
$\therefore$ correlation i^{th} and j^{th} variables is cosine of angle between $\mathbf{x}_{(i)}^{*}$ and $\mathbf{x}_{(j)}^{*}$

D Generalized Variance

Desire a single value which summarizes variability of multivariate observations.

- Recall: $\mathbf{S}$ is a function of deviation vectors $\mathbf{x}^*_{(1)}, \cdots, \mathbf{x}^*_{(p)}$

- It can be shown that $|\mathbf{S}| = \frac{(\text{volume})^2}{(n-1)^p}$
  where volume is the p-dimensional volume of the p-dimensional "box" formed by $\mathbf{x}^*_{(1)}, \cdots, \mathbf{x}^*_{(p)}$



- $|\mathbf{S}|$ is "generalized sample variance"

- $|\mathbf{S}|$ larger as $\mathbf{x}^*_{(1)}, \cdots, \mathbf{x}^*_{(p)}$ are re-oriented to be nearly perpendicular (without changing lengths)
- $|\mathbf{S}|$ larger when $\mathbf{x}^*_{(i)}$ is increased in length ($\mathbf{x}_{(i)}$ multiplied by $c > 1$) without changing orientation
- $|\mathbf{S}| \cong 0$ when any $\mathbf{x}^*_{(i)} \cong \mathbf{0}$ (i.e., small s_{ii})
- $|\mathbf{S}| \cong 0$ when any $\mathbf{x}^*_{(i)}$ lies nearly in $(p-1)$-dim. hyper-plane formed by other deviation vectors

$$\mathbf{x}^*_{(i)} \cong a_1\mathbf{x}^*_{(1)} + \cdots + a_{i-1}\mathbf{x}^*_{(i-1)} + a_{i+1}\mathbf{x}^*_{(i+1)} + \cdots + a_p\mathbf{x}^*_{(p)}$$

- $|\mathbf{S}| = 0$ if one or more of observed variables is a linear function (sum, difference, etc.) of one or more other observed variables
- $|\mathbf{R}| = \frac{(\text{volume})^2}{(n-1)^p}$ where volume is formed by standardized deviation vectors $\frac{\mathbf{x}^*_{(1)}}{\sqrt{s_{11}}}, \cdots, \frac{\mathbf{x}^*_p}{\sqrt{s_{pp}}}$
- $|\mathbf{R}| = \frac{1}{(s_{11}, s_{22}, \cdots, s_{pp})}|\mathbf{S}|$

- $|\mathbf{R}|$ unaffected by multiplying $\mathbf{x}_{(i)}$ by $c \neq 0$
- If $s_{(ij)} = 0$ for all $i \neq j$, $|\mathbf{R}| = 1$
- Alternative to $|\mathbf{S}|$ = "generalized sample variance" is $\text{tr}\{\mathbf{S}\}$ = "<u>total sample variance</u>"

  Recall that $|\mathbf{S}| = \prod_{i=1}^{p} \lambda_i$

$$\text{tr}\{\mathbf{S}\} = \sum_{\mathbf{i=1}}^{\mathbf{p}} \lambda_{\mathbf{i}}$$

  $\text{tr}\{\mathbf{S}\}$ incorporates no multivariate (correlation structure) information