- $\mathbf{AV} = \mathbf{U\Lambda}$

  $\mathbf{A'U} = \mathbf{V\Lambda}$
- $\mathbf{AA'} = \mathbf{U\Lambda^2U'}$

  $\Rightarrow$ squared singular values of $\mathbf{A}$ are eigenvalues of $\mathbf{AA'}$ and columns of $\mathbf{U}$ are eigenvectors of $\mathbf{AA'}$

  $\mathbf{A'A} = \mathbf{V\Lambda^2V'}$

  $\Rightarrow$ squared singular values of $\mathbf{A}$ are eigenvalues of $\mathbf{A'A}$ and columns of $\mathbf{V}$ are eigenvectors of $\mathbf{A'A}$

➢ quadratic form

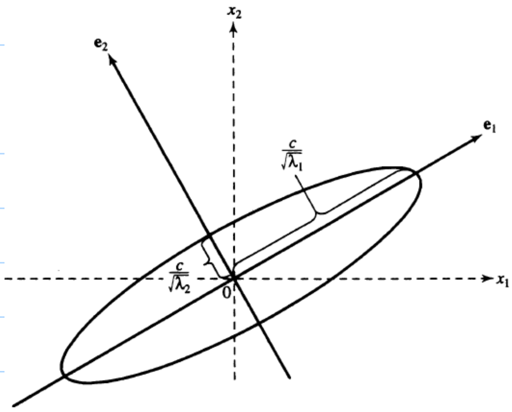
Definition 2A.32. A *quadratic form* $Q(\mathbf{x})$ in the k variables $x_1, x_2, \ldots, x_k$ is $Q(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$, where $\mathbf{x}' = [x_1, x_2, \ldots, x_k]$ and $\mathbf{A}$ is a $k \times k$ symmetric matrix.

$\mathbf{x}'\mathbf{A}\mathbf{x}$

$$= \mathbf{x}'(\lambda_1 \; \mathbf{e}_1 \quad \mathbf{e}_1' \; + \lambda_2 \; \mathbf{e}_2 \quad \mathbf{e}_2' \; + \cdots + \lambda_k \; \mathbf{e}_k \quad \mathbf{e}_k')\,\mathbf{x}$$

$$= \lambda_1(\mathbf{x}'\mathbf{e}_1)(\mathbf{e}_1'\mathbf{x}) + \cdots + \lambda_k(\mathbf{x}'\mathbf{e}_k)(\mathbf{e}_k'\mathbf{x})$$

$$= \lambda_1(\mathbf{x}'\mathbf{e}_1)^2 + \cdots + \lambda_p(\mathbf{x}'\mathbf{e}_p)^2$$



➢ nonnegative definite and positive definite matrix

- When a $k \times k$ symmetric matrix $\mathbf{A}$ is such that

  $$0 \le \mathbf{x}'\mathbf{A}\mathbf{x}$$

  for all $\mathbf{x}' = [x_1, x_2, \ldots, x_k]$, both the matrix $\mathbf{A}$ and the quadratic form are said to be *nonnegative definite*.
- When a $k \times k$ symmetric matrix $\mathbf{A}$ is such that

  $$0 < \mathbf{x}'\mathbf{A}\mathbf{x}$$

  for all vectors $\mathbf{x} \neq \mathbf{0}$, $\mathbf{A}$ or the quadratic form is said to be *positive definite*.
- $\mathbf{A}$ is a positive definite matrix if and only if every eigenvalue of $\mathbf{A}$ is positive

  $\mathbf{A}$ is a nonnegative definite matrix if and only if all of its eigenvalues are greater than or equal to zero
- For nonnegative definite or positive definite matrix

  $\mathbf{x}'\mathbf{A}\mathbf{x} =$
- statistical distance and positive definite matrix

$$(\text{distance})^2 = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{pp}x_p^2$$
$$+ 2(a_{12}x_1x_2 + a_{13}x_1x_3 + \cdots + a_{p-1,p}x_{p-1}x_p)$$
$$= [x_1, x_2, \ldots, x_p] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = \mathbf{x}'\mathbf{A}\mathbf{x}$$

$\Rightarrow$ distance is determined from a positive definite quadratic form $\mathbf{x}'\mathbf{A}\mathbf{x}$

Comment. Let the square of the distance from the point $\mathbf{x}' = [x_1, x_2, \ldots, x_p]$ to the origin be given by $\mathbf{x}'\mathbf{A}\,\mathbf{x}$, where $\mathbf{A}$ is a $p \times p$ symmetric positive definite matrix. Then the square of the distance from $\mathbf{x}$ to an arbitrary fixed point $\boldsymbol{\mu}' = [\mu_1, \mu_2, \ldots, \mu_p]$ is given by the general expression $(\mathbf{x} - \boldsymbol{\mu})'\mathbf{A}(\mathbf{x} - \boldsymbol{\mu})$.

If $p > 2$, the points $\mathbf{x}' = [x_1, x_2, \ldots, x_p]$ a constant distance $c = \sqrt{\mathbf{x}'\mathbf{A}\mathbf{x}}$ from the origin lie on hyperellipsoids $c^2 = \lambda_1(\mathbf{x}'\mathbf{e}_1)^2 + \cdots + \lambda_p(\mathbf{x}'\mathbf{e}_p)^2$, whose axes are given by the eigenvectors of $\mathbf{A}$. The half-length in the direction $\mathbf{e}_i$ is equal to $c/\sqrt{\lambda_i}$, $i = 1, 2, \ldots, p$, where $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the eigenvalues of $\mathbf{A}$.

- square-root matrix
  - Recall: matrix representation of spectral decomposition

Let $\mathbf{A}$ be a $k \times k$ positive definite matrix with the spectral decomposition $\mathbf{A} = \sum_{i=1}^{k} \lambda_i \mathbf{e}_i \mathbf{e}_i'$. Let the normalized eigenvectors be the columns of another matrix

$$\mathbf{P} = [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k]$$

Then

$$\underset{(k\times k)}{\mathbf{A}} = \sum_{i=1}^{k} \lambda_i \underset{(k\times 1)}{\mathbf{e}_i} \underset{(1\times k)}{\mathbf{e}_i'} = \underset{(k\times k)}{\mathbf{P}} \underset{(k\times k)}{\boldsymbol{\Lambda}} \underset{(k\times k)}{\mathbf{P}'}$$

where $\mathbf{P}\mathbf{P}' = \mathbf{P}'\mathbf{P} = \mathbf{I}$ and $\underset{(k\times k)}{\boldsymbol{\Lambda}} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix}$ with $\lambda_i > 0$

- inverse

$$\mathbf{A}^{-1} = \mathbf{P}\boldsymbol{\Lambda}^{-1}\mathbf{P}' = \sum_{i=1}^{k} \frac{1}{\lambda_i} \mathbf{e}_i \mathbf{e}_i'$$

- square-root matrix

$$\mathbf{A}^{1/2} = \sum_{i=1}^{k} \sqrt{\lambda_i}\, \mathbf{e}_i \mathbf{e}_i' = \mathbf{P}\boldsymbol{\Lambda}^{1/2}\mathbf{P}'$$

  - some properties

1. $(\mathbf{A}^{1/2})' = \mathbf{A}^{1/2}$ (that is, $\mathbf{A}^{1/2}$ is symmetric).

2. $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$.

3. $(\mathbf{A}^{1/2})^{-1} = \sum_{i=1}^{k} \frac{1}{\sqrt{\lambda_i}} \mathbf{e}_i \mathbf{e}_i' = \mathbf{P}\boldsymbol{\Lambda}^{-1/2}\mathbf{P}'$, where $\boldsymbol{\Lambda}^{-1/2}$ is a diagonal matrix with $1/\sqrt{\lambda_i}$ as the ith diagonal element.

4. $\mathbf{A}^{1/2}\mathbf{A}^{-1/2} = \mathbf{A}^{-1/2}\mathbf{A}^{1/2} = \mathbf{I}$, and $\mathbf{A}^{-1/2}\mathbf{A}^{-1/2} = \mathbf{A}^{-1}$, where $\mathbf{A}^{-1/2} = (\mathbf{A}^{1/2})^{-1}$.

- matrix inequalities and maximization
  - **Cauchy–Schwarz Inequality.** Let $\mathbf{b}$ and $\mathbf{d}$ be *any* two $p \times 1$ vectors. Then

$$(\mathbf{b}'\mathbf{d})^2 \le (\mathbf{b}'\mathbf{b})(\mathbf{d}'\mathbf{d})$$

with equality if and only if $\mathbf{b} = c\,\mathbf{d}$ (or $\mathbf{d} = c\,\mathbf{b}$) for some constant c.

Extended Cauchy–Schwarz Inequality. Let $\underset{(p\times 1)}{\mathbf{b}}$ and $\underset{(p\times 1)}{\mathbf{d}}$ be any two vectors, and let $\underset{(p\times p)}{\mathbf{B}}$ be a positive definite matrix. Then

$$(\mathbf{b}'\mathbf{d})^2 \le (\mathbf{b}'\mathbf{B}\,\mathbf{b})(\mathbf{d}'\mathbf{B}^{-1}\mathbf{d})$$

with equality if and only if $\mathbf{b} = c\,\mathbf{B}^{-1}\mathbf{d}$ (or $\mathbf{d} = c\,\mathbf{B}\,\mathbf{b}$) for some constant c.

- **Maximization Lemma.** Let $\underset{(p\times p)}{\mathbf{B}}$ be positive definite and $\underset{(p\times 1)}{\mathbf{d}}$ be a given vector. Then, for an arbitrary nonzero vector $\underset{(p\times 1)}{\mathbf{x}}$,

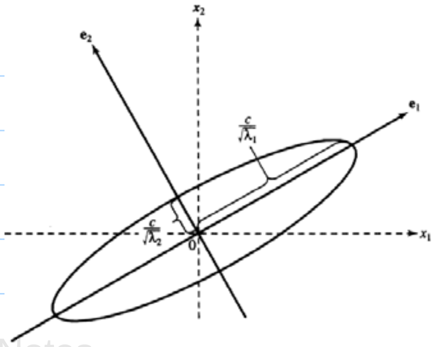
$$\max_{\mathbf{x}\neq\mathbf{0}} \frac{(\mathbf{x}'\mathbf{d})^2}{\mathbf{x}'\mathbf{B}\mathbf{x}} = \mathbf{d}'\mathbf{B}^{-1}\mathbf{d}$$

with the maximum attained when $\underset{(p\times 1)}{\mathbf{x}} = c\underset{(p\times p)(p\times 1)}{\mathbf{B}^{-1}\ \mathbf{d}}$ for any constant $c \neq 0$.

- **Maximization of Quadratic Forms for Points on the Unit Sphere.** Let $\underset{(p\times p)}{\mathbf{B}}$ be a positive definite matrix with eigenvalues $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_p \ge 0$ and associated normalized eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p$. Then

$$\max_{\mathbf{x}\neq\mathbf{0}} \frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \lambda_1 \qquad (\text{attained when } \mathbf{x} = \mathbf{e}_1)$$

$$\min_{\mathbf{x}\neq\mathbf{0}} \frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \lambda_p \qquad (\text{attained when } \mathbf{x} = \mathbf{e}_p)$$



Moreover,

$$\max_{\mathbf{x}\perp \mathbf{e}_1,\ldots,\mathbf{e}_k} \frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \lambda_{k+1} \qquad (\text{attained when } \mathbf{x} = \mathbf{e}_{k+1}, k = 1, 2, \ldots, p-1)$$

where the symbol $\perp$ is read "is perpendicular to."

- approximation of a rectangular matrix by a lower-dimensional matrix
  - sum of squared differences

$$\sum_{i=1}^{m}\sum_{j=1}^{k} (a_{ij} - b_{ij})^2 = \operatorname{tr}[(\mathbf{A}-\mathbf{B})(\mathbf{A}-\mathbf{B})']$$

  - **Result 2A.16.** Let $\mathbf{A}$ be an $m \times k$ matrix of real numbers with $m \ge k$ and singular value decomposition $\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}'$. Let $s < k = \operatorname{rank}(\mathbf{A})$. Then

$$\mathbf{B} = \sum_{i=1}^{s} \lambda_i \mathbf{u}_i \mathbf{v}_i'$$

is the rank-s least squares approximation to $\mathbf{A}$. It minimizes

$$\operatorname{tr}[(\mathbf{A}-\mathbf{B})(\mathbf{A}-\mathbf{B})']$$

over all $m \times k$ matrices $\mathbf{B}$ having rank no greater than s. The minimum value, or error of approximation, is $\sum_{i=s+1}^{k} \lambda_i^2$.

• sample mean, covariance, and correlation as matrix operation

➢ sample mean

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{bmatrix} = \begin{bmatrix} \frac{\mathbf{y}_1'\mathbf{1}}{n} \\ \frac{\mathbf{y}_2'\mathbf{1}}{n} \\ \vdots \\ \frac{\mathbf{y}_p'\mathbf{1}}{n} \end{bmatrix} = \frac{1}{n} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \frac{1}{n}\mathbf{X}'\mathbf{1}$$

➢ sample covariance matrix

$$\mathbf{1}\,\bar{\mathbf{x}}' = \frac{1}{n}\mathbf{1}\mathbf{1}'\mathbf{X} = \begin{bmatrix} \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_p \\ \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_p \\ \vdots & \vdots & \ddots & \vdots \\ \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_p \end{bmatrix}$$

$$\mathbf{X} - \frac{1}{n}\mathbf{1}\mathbf{1}'\mathbf{X} = \begin{bmatrix} x_{11}-\bar{x}_1 & x_{12}-\bar{x}_2 & \cdots & x_{1p}-\bar{x}_p \\ x_{21}-\bar{x}_1 & x_{22}-\bar{x}_2 & \cdots & x_{2p}-\bar{x}_p \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1}-\bar{x}_1 & x_{n2}-\bar{x}_2 & \cdots & x_{np}-\bar{x}_p \end{bmatrix}$$

$$n\,\mathbf{S}_n = \left(\mathbf{X} - \frac{1}{n}\mathbf{1}\mathbf{1}'\mathbf{X}\right)'\left(\mathbf{X} - \frac{1}{n}\mathbf{1}\mathbf{1}'\mathbf{X}\right) = \mathbf{X}'\left(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'\right)\mathbf{X}$$

since $\left(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'\right)'\left(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'\right) = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}' - \frac{1}{n}\mathbf{1}\mathbf{1}' + \frac{1}{n^2}\mathbf{1}\mathbf{1}'\mathbf{1}\mathbf{1}' = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'$

➢ sample correlation matrix

$$\underset{(p\times p)}{\mathbf{D}^{1/2}} = \begin{bmatrix} \sqrt{s_{11}} & 0 & \cdots & 0 \\ 0 & \sqrt{s_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{s_{pp}} \end{bmatrix} \quad \underset{(p\times p)}{\mathbf{D}^{-1/2}} = \begin{bmatrix} \frac{1}{\sqrt{s_{11}}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{s_{22}}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{s_{pp}}} \end{bmatrix}$$

Since

$$\mathbf{S}_n = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{1p} & s_{2p} & \cdots & s_{pp} \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} \frac{s_{11}}{\sqrt{s_{11}}\sqrt{s_{11}}} & \frac{s_{12}}{\sqrt{s_{11}}\sqrt{s_{22}}} & \cdots & \frac{s_{1p}}{\sqrt{s_{11}}\sqrt{s_{pp}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{s_{1p}}{\sqrt{s_{11}}\sqrt{s_{pp}}} & \frac{s_{2p}}{\sqrt{s_{22}}\sqrt{s_{pp}}} & \cdots & \frac{s_{pp}}{\sqrt{s_{pp}}\sqrt{s_{pp}}} \end{bmatrix} = \mathbf{D}^{-1/2}\,\mathbf{S}_n\mathbf{D}^{-1/2}$$

❖ **Reading**: Textbook, 2.2, 2.3, 2.4, 2.7, Supplement 2A, 3.5