(1) Linear Algebra: a brief overview (II)

1. Singular value decomposition

   Let $\mathbf{A}$ be $n \times p$ with rank r. It can be decomposed as

   $$\mathbf{A} = \mathbf{\Gamma_{n\times r}\Lambda_{r\times r}\Delta'_{r\times p}}$$

   where both $\mathbf{\Gamma}$ and $\mathbf{\Delta}$ are column orthonormal. Further $\mathbf{\Lambda} = diag(\sqrt{\lambda_1}, ..., \sqrt{\lambda_r})$ where $\lambda_i > 0, i = 1, ..., r$ are the nonzero eigenvalues of the matrices $\mathbf{A'A}$ and $\mathbf{AA'}$. $\mathbf{\Gamma}$ and $\mathbf{\Delta}$ consist of the corresponding r eigenvectors of these matrices.

2. Quadratic Forms $Q = \mathbf{x'Ax}$ ($\mathbf{A_{p\times p}}$ is symmetric)

   - $\mathbf{A}$ is Positive definite: $Q > 0$ for all $\mathbf{x} \neq \mathbf{0}$
   - $\mathbf{A}$ is Positive semi-definite: $Q \geq 0$ for all $\mathbf{x} \neq \mathbf{0}$
   - If $\mathbf{A}$ is p.d.,

     (a) $\mathbf{A}$ is invertible

     (b) $|\mathbf{A}| > 0$

     (c) All eigenvalues are positive (> 0)

3. A maximization theorem

If $\mathbf{A}_{p\times p}$ and $\mathbf{B}_{p\times p}$ are symmetric and $\mathbf{B}$ is p.d., then for an arbitrary non-zero vector $\mathbf{x}$, then

$$\max_{\mathbf{x}} \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{B}\mathbf{x}} = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p = \min_{\mathbf{x}} \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{B}\mathbf{x}}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$ are the ordered eigenvalues of $\mathbf{B}^{-1}\mathbf{A}$. The vector which maximises (minimises) $\frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{B}\mathbf{x}}$ is the eigenvector of $\mathbf{B}^{-1}\mathbf{A}$ which corresponds to the largest (smallest) eigenvalue of $\mathbf{B}^{-1}\mathbf{A}$.

4. Idempotent Matrices ($\mathbf{A}^2 = \mathbf{A}$)

- All idempotent matrices (except $\mathbf{I}$) are singular
- $\text{rank}(\mathbf{A}) = \text{tr}\mathbf{A}$
- Eigenvalues of idempotent matrices are either 0 or 1.

  Proof:

  Let $\lambda, \mathbf{x}$ be a pair of eigenvalue and eigenvector

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

$$\Rightarrow \mathbf{A}^2\mathbf{x} \;=\; \mathbf{A}(\mathbf{A}\mathbf{x}) \;=\; \mathbf{A}(\lambda\mathbf{x}) \;=\; \lambda\mathbf{A}\mathbf{x} \;=\; \lambda^2\mathbf{x}$$

  But $\quad \mathbf{A}^2\mathbf{x} \;=\; \mathbf{A}\mathbf{x} \;=\; \lambda\mathbf{x}$

  So,

$$\lambda^2\mathbf{x} \;=\; \lambda\mathbf{x} \;\Rightarrow\; \lambda(\lambda - 1)\mathbf{x} \;=\; 0$$

  Since $\mathbf{x} \neq 0$, hence $\lambda = 0 \;\; or \;\; 1$

- Let $\mathbf{A}$ be a symmetric matrix. If all eigenvalues are 1 or 0, $\mathbf{A}$ is idempotent

  Proof:

  If $\mathbf{A}$ is symmetric, there exists an orthogonal matrix $\mathbf{P}$ such that

$$\mathbf{P}'\mathbf{A}\mathbf{P} \;=\; \mathbf{D}$$

  where $\mathbf{D}$ is a diagonal matrix with eigenvalues of $\mathbf{A}$ on the diagonal

  So, $\quad \mathbf{P}'\mathbf{A}\mathbf{P}\mathbf{P}'\mathbf{A}\mathbf{P} = \mathbf{D}^2$

  But, $\quad \mathbf{P}'\mathbf{A}\mathbf{P}\mathbf{P}'\mathbf{A}\mathbf{P} = \mathbf{P}'\mathbf{A}\mathbf{A}\mathbf{P}$

  However if all eigenvalues are 1 or 0, $\mathbf{D} = \mathbf{D}^2$

  $\Rightarrow \quad$ **P'AP** $=$ **P'AAP**

  $\Rightarrow \quad \mathbf{A} = \mathbf{A}\mathbf{A}$

  $\Rightarrow \quad \mathbf{A}$ is idempotent

5. Derivatives

Let $\mathbf{x} = (x_1, x_2, ..., x_p)'$ be a $p \times 1$ vector and assume that $f : \mathbb{R}^p \to \mathbb{R}$ is a scalar function of $\mathbf{x}$. Further, let $\mathbf{a}$ be a $p \times 1$ vector and $\mathbf{A}$ be a $p \times p$ symmetric matrix.

(a)

$$\frac{\partial \mathbf{a}'\mathbf{x}}{\partial \mathbf{x}} = \frac{\partial \mathbf{x}'\mathbf{a}}{\partial \mathbf{x}} = \mathbf{a}$$

(b)

$$\frac{\partial \mathbf{x}'\mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x}$$

(c)

$$\frac{\partial^2 \mathbf{x}'\mathbf{A}\mathbf{x}}{\partial \mathbf{x}\, \partial \mathbf{x}'} = 2\mathbf{A}$$

6. Descriptive statistics

- Data ($\mathbf{X}$)

  Example 3.2 (textbook) A textile shop manager is studying the sales of pullovers over ten different periods. He observes the number of pullovers sold (X_1), variation in price (X_2, in EUR), the advertisement costs in local newspaper (X_3, in EUR) and the presence of a sales assistant (X_4, in hours per period).

$$\mathbf{X} = \begin{pmatrix} 230 & 125 & 200 & 109 \\ 181 & 99 & 55 & 107 \\ 165 & 97 & 105 & 98 \\ 150 & 115 & 85 & 71 \\ 97 & 120 & 0 & 82 \\ 192 & 100 & 150 & 103 \\ 181 & 80 & 85 & 111 \\ 189 & 90 & 120 & 93 \\ 172 & 95 & 110 & 86 \\ 170 & 125 & 130 & 78 \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1' \\ \mathbf{x}_2' \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{x}_n' \end{pmatrix}$$

- Mean ($\overline{\mathbf{x}}$)

$$\overline{\mathbf{x}} = \begin{pmatrix} \overline{x}_1 \\ \cdot \\ \cdot \\ \cdot \\ \overline{x}_p \end{pmatrix} = n^{-1}\mathbf{X}'\mathbf{1_n}$$

  With reference to the pullover data set,

$$\overline{\mathbf{x}} = \begin{pmatrix} 172.7 \\ 104.6 \\ 104 \\ 93.8 \end{pmatrix}$$

- Sample covariance matrix (unbiased: $\mathbf{S}$)

$$\mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})' = \frac{1}{n-1}\mathbf{X}'\mathbf{H}\mathbf{X}$$

where $\mathbf{H} = \mathbf{I}_n - n^{-1}\mathbf{1}_n\mathbf{1}_n'$.

With reference to the pullover data set,

$$\mathbf{S} = \begin{pmatrix} 1152.5 & -88.9 & 1589.7 & 301.6 \\ -88.9 & 244.3 & 102.3 & -101.8 \\ 1589.7 & 102.3 & 2915.6 & 233.7 \\ 301.6 & -101.8 & 233.7 & 197.1 \end{pmatrix}$$

- Correlation ($\mathbf{R}$)

$$\mathbf{R} = \mathbf{D}^{-1/2}\mathbf{S}\mathbf{D}^{-1/2}$$

where $\mathbf{D}$ is a diagonal matrix with form by using the diagonal elements of $\mathbf{S}$.

With reference to the pullover data set,

$$\mathbf{R} = \begin{pmatrix} 1 & -0.17 & 0.87 & 0.63 \\ -0.17 & 1 & 0.12 & -0.46 \\ 0.87 & 0.12 & 1 & 0.31 \\ 0.63 & -0.46 & 0.31 & 1 \end{pmatrix}$$