L05 Sample and Statistics

1. Sample and statistics

   (1) Sample
   Let $X_i \in R^p$, $i = 1, ..., n$, be a random sample from a population with parameters (μ, Σ). Then $X = (X_1, ..., X_n) \in R^{p\times n}$ is a data matrix representing the sample.

   (2) Statistics

   (i) Sum

   $$\sum_{i=1}^n X_i = (X_1, ..., X_n)\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} = X1_n \in R^p \text{ is the sum of all observations.}$$

   (ii) Sample mean
   $\overline{X} = \frac{\sum X_i}{n} = \frac{X1_n}{n} = X(1_n^+)' \in R^p$ is the sample mean.
   For a matrix M with full column rank, $M^+ = (M'M)^{-1}M'$. So $1_n^+ = (1_n'1_n)^{-1}1_n' = \frac{1_n'}{n}$.

   (iii) SSCP (Sum of Squares and Cross Product) matrix
   $\sum_{j=1}^n X_jX_j' = (X_1, ..., X_n)(X_1, ..., X_n)' = XX' \in R^{p\times p}$ is the SSCP matrix.
   Let $X = (x_{ij})_{p\times n}$ and $A = XX' = (a_{ij})_{p\times p}$. Then $a_{ii} = \sum_{j=1}^n x_{ij}^2$ (Sum of squares) and $a_{ij} = \sum_{k=1}^n x_{ik}x_{jk}$ (Sum of Cross-Products).

   (iv) CSSCP (Corrected Sum of Squares and Cross Products) matrix

   $$\begin{aligned}\sum_{j=1}^n (X_j - \overline{X})(X_j - \overline{X})' &= \sum_{j=1}^n X_jX_j' - \frac{1}{n}\left(\sum_j X_j\right)\left(\sum_j X_j\right)' \\ &= XX' - \frac{X1_n1_n'X'}{n} = X(I_n - 1_n1_n^+)X' \in R^{p\times p}.\end{aligned}$$

   is CSSCP matrix. CSSCP is SSCP for $X - \overline{X}1_n'$.

   (v) Sample covariance matrix
   $S = \frac{\text{CSSCP}}{n-1}$ is the sample variance-covariance matrix or simply sample covariance matrix.

   (vi) Sample correlation matrix
   Let $S = (s_{ij})_{p\times p}$ where s_{ii} is also denoted by s_i^2. Then $R = \left(\frac{s_{ij}}{s_i\, s_j}\right)_{p\times p}$ is called the sample correlation matrix.

   Ex1: $R = [\text{diag}(S)]^{-1/2}S[\text{diag}(S)]^{-1/2}$. Let CSSCP $= (a_{ij})_{p\times p}$. Then $R = \left(\frac{a_{ij}}{\sqrt{a_{ii}\ a_{jj}}}\right)_{p\times p}$.
   The diagonal elements of R are 1s.

2. SAS for computing statistics

   (1) Entering sample

$X = \begin{pmatrix}4 & -1 & 3\\ 1 & 3 & 5\end{pmatrix}$	`data a;`
	`input x1 x2;`
	`datalines;`
		`4 1`
		`-1 3`
		`3 5`
		`;`

   (2) Requesting statistics

n, $\overline{X}$ and R	n, $\overline{X}$ and S	n, $\overline{X}$, R, SSCP, CSSCP S

`proc corr;`
`var x1 x2 x3 x4;`
`run;`
	`proc corr nocorr COV;`
	`var x1 x2 x3 x4;`
	`run;`
		`proc corr SSCP CSSCP COV;`
		`var x1 x2 x3 x4;`
		`run;`

3. SAS for principal component analysis

(1) Enter Σ into SAS

$\Sigma = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix}$	`data a (type='cov'); _TYPE_='COV'; input _NAME_ $ x1 x2 x3; datalines; x1 1 -2 0 x2 -2 5 0 x3 0 0 2 ;`	`x1 1 -2 0 x2 . 5 0 x3 . . 2` `x1 1 . . x2 -2 5 . x3 0 0 2`

```
data a (type='cov');
    _TYPE_='COV';
    input _NAME_ $ x1 x2 x3;
    datalines;
    x1  1 -2 0
    x2 -2  5 0
    x3  0  0 2
    ;
```

```
x1 1 -2 0
x2 .  5 0
x3 .  . 2

x1 1  .  .
x2 -2 5  .
x3 0  0  2
```

(2) Eigenvalue decomposition for Σ

```
proc princomp cov;
     var x1 x2 x3;
     run;
```

Total Variables: 3

Total Variances: $\lambda_1 + \lambda_2 + \lambda_3$

EVs of Covariance Matrix

EVs	Difference	Proportion	Cumulative
λ_1	$\lambda_1 - \lambda_2$	$\frac{\lambda_1}{\lambda_1+\lambda_2+\lambda_3}$	$\frac{\lambda_1}{\lambda_1+\lambda_2+\lambda_3}$
λ_2	$\lambda_2 - \lambda_3$	$\frac{\lambda_2}{\lambda_1+\lambda_2+\lambda_3}$	$\frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+\lambda_3}$
λ_3		$\frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3}$	$\frac{\lambda_1+\lambda_2+\lambda_3}{\lambda_1+\lambda_2+\lambda_3}$

Eigenvectors

	Prin1	Prin2	Prin3
x1	p_{11}	p_{12}	p_{13}
x2	p_{21}	p_{22}	p_{23}
x3	p_{31}	p_{32}	p_{33}

(3) Principal components for standardized X
Convert X with $\mathrm{Cov}(X)\Sigma$, to $Z = V^{-1/2}Z = [\mathrm{diag}(\Sigma)]^{-1/2}X$ called standardized X.
Then $\mathrm{Cov}(Z) = V^{-1/2}\Sigma V^{-1/2} = \rho(X)$, the correlation matrix for X. By EVD of ρ one can find principal components for standardized X.

```
proc princomp;
     var x1 x2 x3;
     run;
```

Total Variables: 3

EVs of Correlation Matrix

EVs	Difference	Proportion	Cumulative
λ_1	$\lambda_1 - \lambda_2$	$\frac{\lambda_1}{3}$	$\frac{\lambda_1}{3}$
λ_2	$\lambda_2 - \lambda_3$	$\frac{\lambda_2}{3}$	$\frac{\lambda_1+\lambda_2}{3}$
λ_3		$\frac{\lambda_3}{3}$	$\frac{\lambda_1+\lambda_2+\lambda_3}{3}$

Eigenvectors

	Prin1	Prin2	Prin3
x1	p_{11}	p_{12}	p_{13}
x2	p_{21}	p_{22}	p_{23}
x3	p_{31}	p_{32}	p_{33}

L06: Parameters of sample and statistics

1. Two operations

   (1) Vectorization

   For matrix $X = (X_1, ..., X_n) \in R^{m\times n}$, $\text{vec}(X) \overset{def}{=\!=} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \in R^{mn}$ is the vectorization of X.

   Vectorization $\text{vec}(\cdot)$ is a linear operation, i.e., $\text{vec}(\alpha A + \beta B) = \alpha \text{vec}(A) + \beta \text{vec}(B)$.
   This one-to-one mapping between $R^{m\times n}$ and R^{mn} preserves the inner product

   $$\langle A,\, B\rangle = \text{tr}(B'A) = [\text{vec}(B)]'[\text{vec}(A)] = \langle \text{vec}(A),\, \text{vec}(B)\rangle$$

   With this one-to-one mapping we see that the Hilbert space $R^{m\times n}$ and R^{mn} are isomorphic.

   (2) Kronecker product

   For $A = (a_{ij})_{p\times q}$ and $B \in R^{m\times n}$, $A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1q}B \\ \vdots & \ddots & \vdots \\ a_{p1}B & \cdots & a_{pq}B \end{pmatrix} \in R^{pm\times qn}$ is the Kronecker product of A and B.

   Keronecker product $\otimes$ is associative, i.e., $(A \otimes B) \otimes C = A \otimes (B \otimes C) = A \otimes B \otimes C$.
   But it is not commutative, i.e., $A \otimes B \neq B \otimes A$.
   However, with column vectors x and y, $x \otimes y' = y' \otimes x = xy'$.

   (3) A formula

   Formula: $\text{vec}(AXB) = (B' \otimes A)\, \text{vec}(X)$ links the two operations.
   It shows that the linear transformation $Y = AXB$ from $X \in R^{p\times q}$ to $Y \in R^{m\times n}$, is rendered as the linear transformation $\text{vec}(Y) = (B'A)\, \text{vec}(X)$ from $\text{vec}(X) \in R^{pq}$ to $\text{vec}(Y) \in R^{mn}$.

   Ex1: For column vectors $x \in R^p$ and $y \in R^q$,

   $$\text{vec}(x \otimes y') = \text{vec}(y' \otimes x) = \text{vec}(xy') = \text{vec}\left[\begin{pmatrix} x_1y_1 & \cdots & x_1y_q \\ \vdots & \ddots & \vdots \\ x_py_1 & \cdots & x_py_q \end{pmatrix}\right] = \begin{pmatrix} y_1x \\ \vdots \\ y_qx \end{pmatrix} = y \otimes x.$$

2. One notation

   (1) Definition

   Suppose $X = (X_1, ..., X_n) \in R^{p\times n}$ is a random matrix such that $E(X) = (\mu_1, ..., \mu_n) = M \in R^{p\times n}$ and $\text{Cov}(X_i,\, X_j) = \begin{cases} \psi_{ii}\Sigma & i = j \\ \psi_{ij}\Sigma & \imath \neq j \end{cases}$. Clearly, $\Sigma \in R^{p\times p}$. Let $\Psi = (\psi_{ij})_{n\times n}$. Then

   $$\text{vec}(X) = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \sim \left(\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}, \begin{pmatrix} \psi_{11}\Sigma & \cdots & \psi_{1n}\Sigma \\ \vdots & \ddots & \vdots \\ \psi_{n1}\Sigma & \cdots & \psi_{nn}\Sigma \end{pmatrix}\right) = (\text{vec}(M),\, \Psi \otimes \Sigma) \overset{\text{denoted}}{=\!=} (M,\, \Sigma,\, \Psi)$$

   Ex2: If $\mu \in R^p$, then $X \sim (\mu,\, \Sigma) = (\mu,\, \Sigma,\, 1)$. If $\beta \in R^p$, then $X \sim (\beta,\, \Sigma,\, c) = (\beta,\, c\Sigma)$.

   (2) Two transformations

   $X \sim (M,\, \Sigma,\, \Psi) \Longrightarrow AXB + T \sim (AMB + T,\, A\Sigma A',\, B'\Psi B)$.
   $X \sim (M,\, \Sigma,\, \Psi) \Longrightarrow X' \sim (M',\, \Psi,\, \Sigma)$.

   Proof. Skipped

   Ex3: $X \in R^p$, $X \sim (\mu,\, \Sigma) = (\mu,\, \Sigma,\, 1)$.
   So $AX + b \sim (A\mu + b,\, A\Sigma A',\, 1) = (A\mu + b,\, A\Sigma A')$.

   (3) Expectations

   $X \sim (M,\, \Sigma,\, \Psi)$, $A \in R^{n\times n} \Longrightarrow E(XAX') = MAM' + \text{tr}(A\Psi)\Sigma$.
   $X \sim (M,\, \Sigma,\, \Psi)$, $B \in R^{p\times p} \Longrightarrow E(X'BX) = M'BM + \text{tr}(B\Sigma)\Psi$.

Proof. We show the second one.

If $X \sim (M, \Sigma, \Psi)$, by the second transformation in (2), $Y = X' \sim (M', \Psi, \Sigma)$.

For $E(X'BX) = E(YBY')$, by the first formula in (3),

$$E(X'BX) = E(YBY') = M'BM + \text{tr}(B\Sigma)\Psi.$$

Ex4: $X \in R^p$ and $X \sim (\mu, \Sigma) = (\mu, \Sigma, 1)$. Then $E(X'AX) = \mu' A\mu + \text{tr}(A\Sigma)$.

3. Parameters of sample and statistics

Suppose $X \in R^{p\times n}$ is a random sample from a population with parameters (μ, Σ).

(1) Sample

$X \in R^{p\times n}$ is a random sample from a population with parameters (μ, Σ). Then $X \sim (\mu 1_n', \Sigma, I_n)$.

Proof. $\text{vec}(X) = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \sim \left(\begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix}, \begin{pmatrix} \Sigma & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Sigma \end{pmatrix} \right) = (\text{vec}(\mu 1_n'), I_n \otimes \Sigma).$

So $X \sim (\mu 1_n', \Sigma, I_n)$.

(2) Sum

$\sum_{i=1}^n X_i \sim (n\mu, n\Sigma)$.

Proof. $X \sim (\mu 1_n', \Sigma, 1) \Longrightarrow \sum_i X_i = X 1_n \sim (\mu 1_n' 1_n, \Sigma, 1_n' I_n 1_n) = (n\mu, \Sigma, n) = (n\mu, n\Sigma)$.

(3) Sample mean

$\overline{X} \sim \left(\mu, \frac{1}{n}\Sigma\right)$

Proof. $X \sim (\mu 1_n', \Sigma, I_n) \Longrightarrow \overline{X} = X\frac{1_n}{n} \sim \left(\mu 1_n' \frac{1_n}{n}, \Sigma, \frac{1_n'}{n} I_n \frac{1_n}{n}\right) = \left(\mu, \Sigma, \frac{1}{n}\right) = \left(\mu, \frac{1}{n}\Sigma\right)$.

(4) $E(\text{CSSCP})$

$E(\text{CSSCP}) = (n-1)\Sigma$

Proof. Under $X \sim (\mu 1_n', \Sigma, I_n)$,

$$\begin{aligned} E(\text{CSSCP}) &= E[X(I_n - 1_n 1_n^+)X'] = (\mu 1_n')(I_n - 1_n 1_n^+)(\mu 1_n')' + [\text{tr}(I_n - 1_n 1_n^+)]\Sigma \\ &= 0 + (n-1)\Sigma = (n-1)\Sigma. \end{aligned}$$

(5) $E(S)$

$E(S) = \Sigma$

Proof. $E(S) = E\left(\frac{\text{CSSCP}}{n-1}\right) = \Sigma$.

Ex5: Sample mean $\overline{X}$ is an UE for the population mean μ since $E(\overline{X}) = \mu$.

Sample covariance matrix S is an UE for the population covariance matrix Σ since $E(S) = \Sigma$.