Part I: Linear Models

Chapter I.3

Quadratic Forms

Topics

To be discussed...

- Definition of quadratic forms
- Basic properties
- Cochran's theorem
- Applications of Cochran's theorem

I.3.1 Definition

Let $\mathbf{Y}$ be a random vector and $A \in \mathbb{R}^{p \times p}$ be a symmetric matrix. Then $\mathbf{Y}'A\mathbf{Y}$ is called **quadratic form**.

I.3.2 Remark

Symmetry of A in Definition I.3.1 is not required, since, from $\mathbf{x}'A\mathbf{x} = \mathbf{x}'A'\mathbf{x}$, we have for any $A \in \mathbb{R}^{p \times p}$

$$\mathbf{x}'A\mathbf{x} = \frac{1}{2}(\mathbf{x}'A\mathbf{x} + \mathbf{x}'A'\mathbf{x}) = \mathbf{x}'A_*\mathbf{x}$$

with the symmetric matrix $A_* = \frac{1}{2}(A + A')$. Therefore, without loss of generality, quadratic forms discussed in the following are based on symmetric matrices.

I.3.3 Lemma

Let $\mathbf{Y}$ be a random vector with $\mathrm{E}\mathbf{Y} = \boldsymbol{\mu}$ and Cov $(\mathbf{Y}) = \Sigma$ and $A \in \mathbb{R}^{p \times p}$. Then,

$$\mathrm{E}\mathbf{Y}'A\mathbf{Y} = \mathrm{trace}(A\Sigma) + \boldsymbol{\mu}'A\boldsymbol{\mu}.$$

Orthogonal projectors & quadratic forms

I.3.4 Lemma

Let $Q \in \mathbb{R}^{p\times p}$ be an orthogonal projector. Then, an eigenvalue λ of Q satisfies $\lambda \in \{0, 1\}$. Furthermore, $\mathrm{rank}(Q) = \mathrm{trace}(Q)$.

I.3.5 Theorem

Let $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, I_p)$ be a random vector and Q be an orthogonal projector. Then,

$$\mathbf{Y}'Q\mathbf{Y} \sim \chi^2\Big(\mathrm{rank}(Q), \frac{1}{2}\boldsymbol{\mu}'Q\boldsymbol{\mu}\Big).$$

Orthogonal projectors & quadratic forms

I.3.4 Lemma

Let $Q \in \mathbb{R}^{p \times p}$ be an orthogonal projector. Then, an eigenvalue λ of Q satisfies $\lambda \in \{0, 1\}$. Furthermore, $\mathrm{rank}(Q) = \mathrm{trace}(Q)$.

Proof: Q orth. projector : $Q^2 = Q$ idempotent, $Q' = Q$

Exer. $\Rightarrow$ eigenvalues of Q are in $\{0, 1\}$.

Since $Q = Q'$, the SVD $Q = V \Lambda V'$ with $\Lambda = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$

$r = \mathrm{rank}(Q)$. $= \mathrm{rank}(\Lambda)$

yields

$$\mathrm{trace}(Q) = \mathrm{trace}(V \Lambda V') = \mathrm{trace}(\Lambda \underbrace{V V'}_{= I_p}) = \mathrm{trace}(\Lambda) = r = \mathrm{rank}(Q)$$

Orthogonal projectors & quadratic forms

I.3.4 Lemma

Let $Q \in \mathbb{R}^{p \times p}$ be an orthogonal projector. Then, an eigenvalue λ of Q satisfies $\lambda \in \{0, 1\}$. Furthermore, $\operatorname{rank}(Q) = \operatorname{trace}(Q)$.

I.3.5 Theorem

Let $\mathbf{Y} \sim N_p(\mu, I_p)$ be a random vector and Q be an orthogonal projector. Then,

$$\mathbf{Y}'Q\mathbf{Y} \sim \chi^2\left(\operatorname{rank}(Q), \frac{1}{2}\mu' Q \mu\right).$$

Proof: let $r = \operatorname{rank}(Q)$, consider SVD $Q = V \Lambda V'$, $\Lambda = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$

Define $X := V'Y \sim N_p(V'\mu, \underbrace{V'V}_{=I_p}) = N_p(V'\mu, I_p)$

Then: $Y'QY = \underbrace{Y'V}_{=X'} \Lambda \underbrace{V'Y}_{=X} = X'\Lambda X = \sum_{i=1}^{r} X_i^2$

$$E X_i = (V'\mu)_i = e_{i,p}' \cdot V'\mu , \qquad e_{i,p} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow \text{i-th pos.}$$

(i-th. vector of standard basis)

$$\sum_{i=1}^{r} (EX_i)^2 = \sum_{i=1}^{r} \left(e_{i,p}' V'\mu \right)^2 = \sum_{i=1}^{r} \mu' V e_{i,p} \cdot e_{i,p}' \cdot V'\mu$$

$$= \mu' V \cdot \underbrace{\left(\sum_{i=1}^{r} e_{i,p} \cdot e_{i,p}' \right)}_{= \Delta} V'\mu$$

$$= \mu' \underbrace{V \Delta V'}_{= Q} \mu = \mu' Q \mu .$$

Cochran's theorem

I.3.6 Theorem (Cochran)*

Let $\mathbf{X} \sim N_p(0, \sigma^2 I_p)$ and $A_1, \ldots, A_n \in \mathbb{R}^{p\times p}_{\geqslant 0}$ be non-negative definite matrices with $\sum_{j=1}^n A_j = I_p$. Let $r_j = \operatorname{rank}(A_j)$, $1 \leqslant j \leqslant n$. Then, the following conditions are equivalent:

1. $\sum_{j=1}^n r_j = p$.
2. $\frac{1}{\sigma^2}\mathbf{X}' A_j \mathbf{X} \sim \chi^2(r_j)$, $1 \leqslant j \leqslant n$
3. $\mathbf{X}' A_j \mathbf{X}$, $1 \leqslant j \leqslant n$, are mutually independent.

I.3.7 Remark

Theorem I.3.6 may be extended in various directions. In Rencher & Schaalje (2008), Theorem 5.6c, a non-central version is presented, that is, one assumes $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \sigma^2 I_p)$ with $\boldsymbol{\mu} \in \mathbb{R}^p$. Then, the following conditions are equivalent:

1. $\sum_{j=1}^n r_j = p$.
2. $\frac{1}{\sigma^2}\mathbf{X}' A_j \mathbf{X} \sim \chi^2(r_j, \boldsymbol{\mu}' A_j \boldsymbol{\mu}/2)$, $1 \leqslant j \leqslant n$
3. $\mathbf{X}' A_j \mathbf{X}$, $1 \leqslant j \leqslant n$, are mutually independent.

*see, e,g., Gut, A. (2009) An Intermediate Course in Probability. 2nd edn., Springer, New York, Section 5.9.

Application of Cochran's theorem

I.3.8 Corollary (see Theorem I.2.15)

Let $p \geqslant 2$, $\mathbf{Z} = (Z_1, \ldots, Z_p)' \sim N_p(\mathbb{0}, \sigma^2 I_p)$, $\overline{Z} = \frac{1}{p}\sum_{j=1}^{p} Z_j$, $S_Z = \frac{1}{p-1}\sum_{j=1}^{p}(Z_j - \overline{Z})^2$, and $E_p = I_p - \frac{1}{p}\mathbb{1}_p\mathbb{1}_p'$. Then:

1. $p\dfrac{\overline{Z}^2}{\sigma^2} \sim \chi^2(1)$ and $\dfrac{1}{\sigma^2}\sum_{j=1}^{p}(Z_j - \overline{Z})^2 = \dfrac{1}{\sigma^2}\mathbf{Z}' E_p \mathbf{Z} \sim \chi^2(p-1)$ are independent.
2. $\dfrac{\overline{Z}^2}{S_Z} \sim F(1, p-1)$

Use Cochran's Theorem:

(1) $E_p + A = I_p$

(2) $\operatorname{rank}(A) = \operatorname{rank}\left(\frac{1}{p}\mathbb{1}_{p\times p}\right) = \operatorname{trace}\left(\frac{1}{p}\mathbb{1}_{p\times p}\right) = p\cdot\frac{1}{p} = 1$

$\operatorname{rank}(E_p) = \operatorname{trace}(E_p) = \operatorname{tr}(I_p - A) = \operatorname{tr}(I_p) - \operatorname{tr}(A)$

$= p - 1$

$\Rightarrow \operatorname{rank}(A) + \operatorname{rank}(E_p) = p = \operatorname{rank}(I_p)$

$\overset{\text{Cochran}}{\Rightarrow}$ $z'Az$, $z'E_p z$ are indep.

and $\frac{z'Az}{\sigma^2} \sim \chi^2(1)$, $\frac{z'E_p z'}{\sigma^2} \sim \chi^2(p-1)$

$\Rightarrow \bar{z}^2/S_z \sim F(1, p-1)$

Application of Cochran's theorem

I.3.8 Corollary (see Theorem I.2.15)

Let $p \geqslant 2$, $Z = (Z_1, \ldots, Z_p)' \sim N_p(0, \sigma^2 I_p)$, $\overline{Z} = \frac{1}{p}\sum_{j=1}^{p} Z_j$, $S_Z = \frac{1}{p-1}\sum_{j=1}^{p}(Z_j - \overline{Z})^2$, and $E_p = I_p - \frac{1}{p}\mathbb{1}_p\mathbb{1}_p'$. Then:

1. $p\dfrac{\overline{Z}^2}{\sigma^2} \sim \chi^2(1)$ and $\dfrac{1}{\sigma^2}\sum_{j=1}^{p}(Z_j - \overline{Z})^2 = \dfrac{1}{\sigma^2}Z'E_pZ \sim \chi^2(p-1)$ are independent.
2. $\dfrac{\overline{Z}^2}{S_Z} \sim F(1, p-1)$

Proof: • $p \cdot \overline{z}^2 = \frac{1}{p} \cdot z'\mathbb{1}_p \cdot \mathbb{1}_p' z = z'\left(\frac{1}{p}\mathbb{1}_{p\times p}\right)z$

• $\sum_{j=1}^{p}(z_j - \overline{z})^2 = z'E_pz$ (Theorem I.2.15)

- $A = \frac{1}{p}\cdot\mathbb{1}_{p\times p}$, $E_p = I_p - A$ orthogonal projectors

• $A = A'$, $A^2 = \frac{1}{p}\mathbb{1}_p\mathbb{1}_p' \cdot \frac{1}{p}\mathbb{1}_p\mathbb{1}_p' = \frac{1}{p^2}\cdot p \cdot \mathbb{1}_{p\times p} = A$

• $E_p = E_p'$, $E_p^2 = (I_p - A)(I_p - A) = I_p - A - A + \underset{=A}{A^2} = I_p - A$

Two-sample case and F-distribution

I.3.9 Corollary

Let $X_1, \ldots, X_{n_1} \overset{\text{iid}}{\sim} N(\mu_1, \sigma_1^2)$, $Y_1, \ldots, Y_{n_2} \overset{\text{iid}}{\sim} N(\mu_2, \sigma_2^2)$ be independent samples of random variables with $n_1, n_2 \geqslant 2$ and

$$\widehat{\sigma}_1^2 = \frac{1}{n_1 - 1} \sum_{j=1}^{n_1} (X_j - \overline{X})^2 = \frac{1}{n_1 - 1} \mathbf{X}' E_{n_1} \mathbf{X} \quad \text{and} \quad \widehat{\sigma}_2^2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (Y_j - \overline{Y})^2 = \frac{1}{n_2 - 1} \mathbf{Y}' E_{n_2} \mathbf{Y}.$$

Then, we get

1. $\mathbf{X}' E_{n_1} \mathbf{X} / \sigma_1^2 \sim \chi^2(n_1 - 1)$, $\mathbf{Y}' E_{n_2} \mathbf{Y} / \sigma_2^2 \sim \chi^2(n_2 - 1)$, and $\boxed{\dfrac{\widehat{\sigma}_1^2 / \sigma_1^2}{\widehat{\sigma}_2^2 / \sigma_2^2} \sim F(n_1 - 1, n_2 - 1)}$
2. For $\sigma_1 = \sigma_2 = \sigma$:
   - the F-statistic $\boxed{F = \dfrac{\widehat{\sigma}_1^2}{\widehat{\sigma}_2^2}}$ is F-distributed with $n_1 - 1$ and $n_2 - 1$ degrees of freedom.
   - $\boxed{\dfrac{(n_1 - 1)\widehat{\sigma}_1^2 + (n_2 - 1)\widehat{\sigma}_2^2}{\sigma^2} \sim \chi^2(n_1 + n_2 - 2).}$

Two-sample case and F-distribution

I.3.9 Corollary

Let $X_1, \dots, X_{n_1} \overset{\text{iid}}{\sim} N(\mu_1, \sigma_1^2)$, $Y_1, \dots, Y_{n_2} \overset{\text{iid}}{\sim} N(\mu_2, \sigma_2^2)$ be independent samples of random variables with $n_1, n_2 \geqslant 2$ and

$$\widehat{\sigma}_1^2 = \frac{1}{n_1 - 1} \sum_{j=1}^{n_1} (X_j - \overline{X})^2 = \frac{1}{n_1 - 1} \mathbf{X}' E_{n_1} \mathbf{X} \quad \text{and} \quad \widehat{\sigma}_2^2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (Y_j - \overline{Y})^2 = \frac{1}{n_2 - 1} \mathbf{Y}' E_{n_2} \mathbf{Y}.$$

$z_j = x_j - \mu_1$
$\bar{z} = \bar{x} - \mu_1$

$z' E_n z$

Then, we get

1. $\mathbf{X}' E_{n_1} \mathbf{X} / \sigma_1^2 \sim \chi^2(n_1 - 1)$, $\mathbf{Y}' E_{n_2} \mathbf{Y} / \sigma_2^2 \sim \chi^2(n_2 - 1)$, and $\boxed{\dfrac{\widehat{\sigma}_1^2 / \sigma_1^2}{\widehat{\sigma}_2^2 / \sigma_2^2} \sim F(n_1 - 1, n_2 - 1)}$

   independent.

2. For $\sigma_1 = \sigma_2 = \sigma$:
   - the F-statistic $\boxed{F = \dfrac{\widehat{\sigma}_1^2}{\widehat{\sigma}_2^2}}$ is F-distributed with $n_1 - 1$ and $n_2 - 1$ degrees of freedom.
   - $\boxed{\dfrac{(n_1 - 1)\widehat{\sigma}_1^2 + (n_2 - 1)\widehat{\sigma}_2^2}{\sigma^2} \sim \chi^2(n_1 + n_2 - 2).}$

An extension to random vectors

I.3.10 Theorem[†]

Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} N_p(\mu, \Sigma)$ with $\mu \in \mathbb{R}^p, \Sigma \in \mathbb{R}^{p \times p}_{>0}$ and $n > p$. Then:

1. $\overline{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim N_p(\mu, \frac{1}{n}\Sigma)$
2. $n\|\overline{X} - \mu\|^2_{\Sigma} \sim \chi^2(n)$.
3. For the **sample covariance matrix** $\hat{\Sigma} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X})(X_i - \overline{X})'$, we have
   - $E\hat{\Sigma} = \Sigma$
   - $(n-1)\hat{\Sigma}$ has a so-called **Wishart**-distribution
   - $n\|\overline{X} - \mu\|^2_{\hat{\Sigma}}$ has an **Hotellings-T^2-distribution** with parameters p and $n-1$ (for short, $n\|\overline{X} - \mu\|^2_{\hat{\Sigma}} \sim T_{p,n-1}$).

     For $m > p$: if $Y \sim T_{p,m}$ then $\frac{m-p+1}{mp} Y \sim F_{p,m-p+1}$

[†] cf. T. W. Anderson (2003) Introduction to Multivariate Statistical Analysis, 3rd ed., New York: Wiley, 2003, Chapters 3 & 5.