Random Vectors[1]

STA 302 Fall 2013

[1] See last slide for copyright information.

Overview

Random Vectors and Matrices

See Chapter 3 of *Linear models in statistics* for more detail.

A *random matrix* is just a matrix of random variables. Their joint probability distribution is the distribution of the random matrix. Random matrices with just one column (say, $p \times 1$) may be called *random vectors*.

Expected Value

The expected value of a matrix is defined as the matrix of expected values. Denoting the $p \times c$ random matrix $\mathbf{X}$ by $[X_{i,j}]$,

$$E(\mathbf{X}) = [E(X_{i,j})].$$

Immediately we have natural properties like

$$
\begin{aligned}
E(\mathbf{X}+\mathbf{Y}) &= E([X_{i,j}]+[Y_{i,j}]) \\
&= [E(X_{i,j}+Y_{i,j})] \\
&= [E(X_{i,j})+E(Y_{i,j})] \\
&= [E(X_{i,j})]+[E(Y_{i,j})] \\
&= E(\mathbf{X})+E(\mathbf{Y}).
\end{aligned}
$$

Moving a constant through the expected value sign

Let $\mathbf{A} = [a_{i,j}]$ be an $r \times p$ matrix of constants, while $\mathbf{X}$ is still a $p \times c$ random matrix. Then

$$\begin{aligned} E(\mathbf{AX}) &= E\left(\left[\sum_{k=1}^{p} a_{i,k}X_{k,j}\right]\right) \\ &= \left[E\left(\sum_{k=1}^{p} a_{i,k}X_{k,j}\right)\right] \\ &= \left[\sum_{k=1}^{p} a_{i,k}E(X_{k,j})\right] \\ &= \mathbf{A}E(\mathbf{X}). \end{aligned}$$

Similar calculations yield $E(\mathbf{AXB}) = \mathbf{A}E(\mathbf{X})\mathbf{B}$.

Variance-Covariance Matrices

Let $\mathbf{X}$ be a $p \times 1$ random vector with $E(\mathbf{X}) = \boldsymbol{\mu}$. The *variance-covariance matrix* of $\mathbf{X}$ (sometimes just called the *covariance matrix*), denoted by $cov(\mathbf{X})$, is defined as

$$cov(\mathbf{X}) = E\left\{(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})'\right\}.$$

$cov(\mathbf{X}) = E\{(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})'\}$

$$
\begin{aligned}
cov(\mathbf{X}) &= E\left\{\begin{pmatrix} X_1-\mu_1 \\ X_2-\mu_2 \\ X_3-\mu_3 \end{pmatrix}\begin{pmatrix} X_1-\mu_1 & X_2-\mu_2 & X_3-\mu_3 \end{pmatrix}\right\} \\
&= E\left\{\begin{pmatrix} (X_1-\mu_1)^2 & (X_1-\mu_1)(X_2-\mu_2) & (X_1-\mu_1)(X_3-\mu_3) \\ (X_2-\mu_2)(X_1-\mu_1) & (X_2-\mu_2)^2 & (X_2-\mu_2)(X_3-\mu_3) \\ (X_3-\mu_3)(X_1-\mu_1) & (X_3-\mu_3)(X_2-\mu_2) & (X_3-\mu_3)^2 \end{pmatrix}\right. \\
&= \left(\begin{matrix} E\{(X_1-\mu_1)^2\} & E\{(X_1-\mu_1)(X_2-\mu_2)\} & E\{(X_1-\mu_1)(X_3 - \\ E\{(X_2-\mu_2)(X_1-\mu_1)\} & E\{(X_2-\mu_2)^2\} & E\{(X_2-\mu_2)(X_3 - \\ E\{(X_3-\mu_3)(X_1-\mu_1)\} & E\{(X_3-\mu_3)(X_2-\mu_2)\} & E\{(X_3-\mu_3)^2\} \end{matrix}\right. \\
&= \begin{pmatrix} Var(X_1) & Cov(X_1,X_2) & Cov(X_1,X_3) \\ Cov(X_1,X_2) & Var(X_2) & Cov(X_2,X_3) \\ Cov(X_1,X_3) & Cov(X_2,X_3) & Var(X_3) \end{pmatrix}.
\end{aligned}
$$

So, the covariance matrix $cov(\mathbf{X})$ is a $p \times p$ symmetric matrix with variances on the main diagonal and covariances on the off-diagonals.

Analogous to $Var(a\,X) = a^2\,Var(X)$

Let $\mathbf{X}$ be a $p \times 1$ random vector with $E(\mathbf{X}) = \boldsymbol{\mu}$ and $cov(\mathbf{X}) = \boldsymbol{\Sigma}$, while $\mathbf{A} = [a_{i,j}]$ is an $r \times p$ matrix of constants. Then

$$\begin{aligned}
cov(\mathbf{AX}) &= E\left\{(\mathbf{AX} - \mathbf{A}\boldsymbol{\mu})(\mathbf{AX} - \mathbf{A}\boldsymbol{\mu})'\right\} \\
&= E\left\{\mathbf{A}(\mathbf{X} - \boldsymbol{\mu})\left(\mathbf{A}(\mathbf{X} - \boldsymbol{\mu})\right)'\right\} \\
&= E\left\{\mathbf{A}(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})'\mathbf{A}'\right\} \\
&= \mathbf{A}E\{(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})'\}\mathbf{A}' \\
&= \mathbf{A}cov(\mathbf{X})\mathbf{A}' \\
&= \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'
\end{aligned}$$

Positive definite is a natural assumption

For covariance matrices

- $cov(\mathbf{X}) = \boldsymbol{\Sigma}$
- $\boldsymbol{\Sigma}$ positive definite means $\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a} > 0.$ for all $\mathbf{a} \neq \mathbf{0}$.
- $Y = \mathbf{a}'\mathbf{X} = a_1 X_1 + \cdots + a_p X_p$ is a scalar random variable.
- $Var(Y) = \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}$
- $\boldsymbol{\Sigma}$ positive definite just says that the variance of any (non-trivial) linear combination is positive.
- This is often what you want (but not always).

Matrix of covariances between two random vectors

Let $\mathbf{X}$ be a $p \times 1$ random vector with $E(\mathbf{X}) = \boldsymbol{\mu}_x$ and let $\mathbf{Y}$ be a $q \times 1$ random vector with $E(\mathbf{Y}) = \boldsymbol{\mu}_y$. The $p \times q$ matrix of covariances between the elements of $\mathbf{X}$ and the elements of $\mathbf{Y}$ is

$$C(\mathbf{X}, \mathbf{Y}) = E\left\{(\mathbf{X} - \boldsymbol{\mu}_x)(\mathbf{Y} - \boldsymbol{\mu}_y)'\right\}.$$

Adding a constant has no effect

On variances and covariances

It's clear from the definitions:

- $cov(\mathbf{X}) = E\left\{(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})'\right\}$
- $C(\mathbf{X},\mathbf{Y}) = E\left\{(\mathbf{X}-\boldsymbol{\mu}_x)(\mathbf{Y}-\boldsymbol{\mu}_y)'\right\}$

That

- $cov(\mathbf{X}+\mathbf{a}) = cov(\mathbf{X})$
- $C(\mathbf{X}+\mathbf{a},\mathbf{Y}+\mathbf{b}) = C(\mathbf{X},\mathbf{Y})$

For example, $E(\mathbf{X}+\mathbf{a}) = \boldsymbol{\mu}+\mathbf{a}$, so

$$\begin{aligned} cov(\mathbf{X}+\mathbf{a}) &= E\left\{(\mathbf{X}+\mathbf{a}-(\boldsymbol{\mu}+\mathbf{a}))(\mathbf{X}+\mathbf{a}-(\boldsymbol{\mu}+\mathbf{a}))'\right\} \\ &= E\left\{(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})'\right\} \\ &= cov(\mathbf{X}) \end{aligned}$$

Moment-generating function

Of a p-dimensional random vector $\mathbf{X}$

- $M_{\mathbf{X}}(\mathbf{t}) = E\left(e^{\mathbf{t}'\mathbf{X}}\right)$
- Corresponds uniquely to the probability distribution.

Section 4.3 of *Linear models in statistics* has some material on moment-generating functions.

$M_{\mathbf{AX}}(\mathbf{t}) = M_{\mathbf{X}}(\mathbf{A}'\mathbf{t})$

Analogue of $M_{aX}(t) = M_X(at)$

$$\begin{aligned} M_{\mathbf{AX}}(\mathbf{t}) &= E\left(e^{\mathbf{t}'\mathbf{AX}}\right) \\ &= E\left(e^{\left(\mathbf{A}'\mathbf{t}\right)'\mathbf{X}}\right) \\ &= M_{\mathbf{X}}(\mathbf{A}'\mathbf{t}) \end{aligned}$$

Note that $\mathbf{t}$ is the same length as $\mathbf{Y} = \mathbf{AX}$: The number of rows in $\mathbf{A}$.

$M_{\mathbf{X}+\mathbf{c}}(\mathbf{t}) = e^{\mathbf{t}'\mathbf{c}} M_{\mathbf{X}}(\mathbf{t})$

Analogue of $M_{X+c}(t) = e^{ct} M_X(t)$

$$
\begin{aligned}
M_{\mathbf{X}+\mathbf{c}}(\mathbf{t}) &= E\left(e^{\mathbf{t}'(\mathbf{X}+\mathbf{c})}\right) \\
&= E\left(e^{\mathbf{t}'\mathbf{X}+\mathbf{t}'\mathbf{c}}\right) \\
&= e^{\mathbf{t}'\mathbf{c}}\, E\left(e^{\mathbf{t}'\mathbf{X}}\right) \\
&= e^{\mathbf{t}'\mathbf{c}} M_{\mathbf{X}}(\mathbf{t})
\end{aligned}
$$

Independence

Two random vectors $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent if and only if the moment-generating function of their joint distribution is the product of their moment-generating functions.

Proof: Suppose $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent, with

$\mathbf{X} = \left(\frac{\mathbf{X}_1}{\mathbf{X}_2} \right)$ and $\mathbf{t} = \left(\frac{\mathbf{t}_1}{\mathbf{t}_2} \right)$. Then

$$
\begin{aligned}
M_{\mathbf{X}}(\mathbf{t}) &= E\left(e^{\mathbf{t}'\mathbf{X}}\right) \\
&= E\left(e^{\mathbf{t}_1'\mathbf{X}_1+\mathbf{t}_2'\mathbf{X}_2}\right) = E\left(e^{\mathbf{t}_1'\mathbf{X}_1}e^{\mathbf{t}_2'\mathbf{X}_2}\right) \\
&= \int\int e^{\mathbf{t}_1'\mathbf{x}_1}e^{\mathbf{t}_2'\mathbf{x}_2} f_{\mathbf{X}_1}(\mathbf{x}_1) f_{\mathbf{X}_2}(\mathbf{x}_2)\, d(\mathbf{x}_1) d(\mathbf{x}_2) \\
&= \int e^{\mathbf{t}_2'\mathbf{x}_2} \left(\int e^{\mathbf{t}_1'\mathbf{x}_1} f_{\mathbf{X}_1}(\mathbf{x}_1)\, d(\mathbf{x}_1) \right) f_{\mathbf{X}_2}(\mathbf{x}_2)\, d(\mathbf{x}_2) \\
&= \int e^{\mathbf{t}_2'\mathbf{x}_2} M_{\mathbf{X}_1}(\mathbf{t}_1) f_{\mathbf{X}_2}(\mathbf{x}_2)\, d(\mathbf{x}_2) \\
&= M_{\mathbf{X}_1}(\mathbf{t}_1) M_{\mathbf{X}_2}(\mathbf{t}_2)
\end{aligned}
$$

By uniqueness, it's an if and only if.

$\mathbf{X}_1$ and $\mathbf{X}_2$ independent implies that $\mathbf{Y}_1 = g_1(\mathbf{X}_1)$ and $\mathbf{Y}_2 = g_2(\mathbf{X}_2)$ are independent.

Let

$$\mathbf{Y} = \left(\frac{\mathbf{Y}_1}{\mathbf{Y}_2}\right) = \left(\frac{g_1(\mathbf{X}_1)}{g_2(\mathbf{X}_2)}\right) \text{ and } \mathbf{t} = \left(\frac{\mathbf{t}_1}{\mathbf{t}_2}\right). \text{ Then}$$

$$\begin{aligned}
M_{\mathbf{Y}}(\mathbf{t}) &= E\left(e^{\mathbf{t}'\mathbf{Y}}\right) \\
&= E\left(e^{\mathbf{t}_1'\mathbf{Y}_1+\mathbf{t}_2'\mathbf{Y}_2}\right) = E\left(e^{\mathbf{t}_1'\mathbf{Y}_1}e^{\mathbf{t}_2'\mathbf{Y}_2}\right) \\
&= E\left(e^{\mathbf{t}_1'g_1(\mathbf{X}_1)}e^{\mathbf{t}_2'g_2(\mathbf{X}_2)}\right) \\
&= \int\int e^{\mathbf{t}_1'g_1(\mathbf{x}_1)}e^{\mathbf{t}_2'g_2(\mathbf{x}_2)} f_{\mathbf{X}_1}(\mathbf{x}_1)f_{\mathbf{X}_2}(\mathbf{x}_2)\, d(\mathbf{x}_1)d(\mathbf{x}_2) \\
&= M_{g_1(\mathbf{X}_1)}(\mathbf{t}_1)M_{g_2(\mathbf{X}_2)}(\mathbf{t}_2) \\
&= M_{\mathbf{Y}_1}(\mathbf{t}_1)M_{\mathbf{Y}_2}(\mathbf{t}_2)
\end{aligned}$$

So $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are independent.

Copyright Information