STAT 830
The Multivariate Normal Distribution

Richard Lockhart

Simon Fraser University

STAT 830 — Fall 2013

What I assume you already know

- The basics of normal distributions in 1 dimension.

The Multivariate Normal Distribution

- **Definition**: $Z \in R^1 \sim N(0,1)$ iff
$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} .$$
- **Definition**: $Z \in R^p \sim MVN(0,I)$ if and only if $Z = (Z_1,\ldots,Z_p)^T$ with the Z_i independent and each $Z_i \sim N(0,1)$.
- In this case according to our theorem
$$f_Z(z_1,\ldots,z_p) = \prod \frac{1}{\sqrt{2\pi}} e^{-z_i^2/2} = (2\pi)^{-p/2} \exp\{-z^T z/2\} ;$$
superscript T denotes matrix transpose.
- **Definition**: $X \in R^p$ has a multivariate normal distribution if it has same distribution as $AZ+\mu$ for some $\mu \in R^p$, some $p \times p$ matrix of constants A and $Z \sim MVN(0,I)$.

The Multivariate Normal Density

- Matrix A singular: X does not have a density.
- A invertible: derive multivariate normal density by change of variables:

$$X = AZ + \mu \Leftrightarrow Z = A^{-1}(X - \mu) \qquad \frac{\partial X}{\partial Z} = A \qquad \frac{\partial Z}{\partial X} = A^{-1}.$$

- So

$$\begin{aligned} f_X(x) &= f_Z(A^{-1}(x-\mu))|\det(A^{-1})| \\ &= \frac{\exp\{-(x-\mu)^T (A^{-1})^T A^{-1}(x-\mu)/2\}}{(2\pi)^{p/2}|\det A|}. \end{aligned}$$

The Multivariate Normal Density continued

- Now define $\Sigma = AA^T$ and notice that

$$\Sigma^{-1} = (A^T)^{-1}A^{-1} = (A^{-1})^T A^{-1}$$

and

$$\det \Sigma = \det A \det A^T = (\det A)^2 \,.$$

- Thus f_X is

$$\frac{\exp\{-(x-\mu)^T\Sigma^{-1}(x-\mu)/2\}}{(2\pi)^{p/2}(\det \Sigma)^{1/2}} \,;$$

the $MVN(\mu, \Sigma)$ density.

- Note density is the same for all A such that $AA^T = \Sigma$.
- This justifies the notation $MVN(\mu, \Sigma)$.

The Multivariate Normal Density continued

- For which μ, Σ is this a density?
- Any μ but if $x \in R^p$ then, putting $y = A^T x$,

$$x^T \Sigma x = x^T A A^T x = (A^T x)^T (A^T x) = \sum_1^p y_i^2 \geq 0$$

- Inequality strict except for $y = 0$ which is equivalent to $x = 0$.
- Thus Σ is a positive definite symmetric matrix.
- Conversely, if Σ is a positive definite symmetric matrix then there is a square invertible matrix A such that $AA^T = \Sigma$ so that there is a $MVN(\mu, \Sigma)$ distribution.
- A can be found via the Cholesky decomposition, e.g.

Singular cases

- When A is singular X will not have a density.
- $\exists a$ such that $P(a^T X = a^T \mu) = 1$
- X is confined to a hyperplane.
- Still true: distribution of X depends only on $\Sigma = AA^T$
- if $AA^T = BB^T$ then $AZ + \mu$ and $BZ + \mu$ have the same distribution.
- Proof by mgfs or characteristic functions.

Equality in distribution

- We say X and Y have the same distribution if, for all A,

$$P(X \in A) = P(Y \in A).$$

- If X has density f then X and Y have the same distribution iff Y has density f.
- If $X \in \mathbb{R}^p$ then the moment generating function (mgf) of X is

$$M_X(t) = \mathrm{E}\left[e^{t^T X}\right]$$

  for $t \in \mathbb{R}^p$.
- If $X \in \mathbb{R}^p$ then the characteristic function (cf) of X is

$$\phi_X(t) = \mathrm{E}\left[e^{it^T X}\right]$$

  for $t \in \mathbb{R}^p$; the symbol i is the imaginary unit, $i^2 = -1$.
- cf is complex number defined for every $t \in \mathbb{R}^p$. The mgf may well be ∞ for any $t \neq 0$.

Equality in distribution 2

- If there is an $\epsilon > 0$ such that
$$M_Y(t) = M_X(t)$$
for all t such that $||t|| = \sqrt{t^T t} < \epsilon$ then X and Y have the same distribution.
- If
$$\phi_Y(t) = \phi_X(t)$$
for all $t \in \mathbb{R}^p$ then X and Y have the same distribution.

Application to MVN

- If Z is $\text{MVN}_p(0, I)$ then

$$\begin{aligned}\phi_Z(t) &= \mathrm{E}\left(\exp\{it^Z\}\right) = \mathrm{E}\left(\exp\{\sum_j it_j Z_j\}\right) \\ &= \mathrm{E}\left(\prod_h \exp\{it_j Z_j\}\right) = \prod_h \mathrm{E}\left(\exp\{it_h Z_h\}\right) \\ &= \prod_j \phi_N(t_j)\end{aligned}$$

where ϕ_N denotes the cf of a $N(0, 1)$ variate.

Application to MVN 2

- The cf of $N(0,1)$ is

$$\begin{aligned}\phi_N(t) &= \mathrm{E}\left(\exp\{itZ\}\right) = \int_{-\infty}^{\infty} \exp\{itz - z^2/2\} dz/\sqrt{2\pi} \\ &= \int_{-\infty}^{\infty} \exp\{-t^2/2 - (z-it)^2/2\} dz/\sqrt{2\pi} \\ &= \exp(-t^2/2) \int_{-\infty}^{\infty} \exp\{-(z-it)^2/2\} dz/\sqrt{2\pi} = \exp(-t^2/2)\end{aligned}$$

- So the multivariate cf above is

$$\phi_Z(t) = \prod_j \exp\{-t_j^2/2\} = \exp\{-t^T t/2\}.$$

- Notice the use of normal density with mean $\mu = it$; works by magic.

General Case

- If $X = AZ + \mu$ with $Z \in \mathbb{R}^q$, A a $p \times q$ matrix and $\mu \in \mathbb{R}^p$ then

$$\mathrm{E}\left(\exp\{it^T X\}\right) = \exp(it^T \mu)\phi_Z(A^T t) = \exp(it^T \mu - t^T A A^T t/2)$$

- Depends only on μ and $\Sigma = AA^T$ so distribution of X depends only on its mean and variance.
- The mgf is

$$M_X(t) == \exp(t^T \mu + t^T A A^T t/2)$$

Properties of the *MVN* distribution

1. All margins are multivariate normal: if

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

and

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

then $X \sim MVN(\mu, \Sigma) \Rightarrow X_1 \sim MVN(\mu_1, \Sigma_{11})$.

2. All conditionals are normal: the conditional distribution of X_1 given $X_2 = x_2$ is $MVN(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$
3. $MX + \nu \sim MVN(M\mu + \nu, M\Sigma M^T)$: affine transformation of MVN is normal.