Chapter 2

Multivariate normal distribution

2.1 Basic facts

Let $Z_1, Z_2, \dots, Z_n$ be independent $N(0,1)$ random variables. When treated as the coordinates of a point in $\mathbb{R}^n$ they define a random vector $\mathbf{Z}$, whose (joint) density function is

$$f(\mathbf{z}) = (2\pi)^{-n/2} \exp\left(-\tfrac{1}{2}\sum\nolimits_i z_i^2\right) = (2\pi)^{-n/2} \exp\left(-\tfrac{1}{2}\|\mathbf{z}\|^2\right).$$

Such a random vector is said to have a ***spherical normal distribution***.

The ***chi-square***, χ_n^2, is defined as the distribution of the sum of squares $R^2 := Z_1^2 + \dots + Z_n^2$ of independent $N(0,1)$ random variables. The ***non-central chi-square***, $\chi_n^2(\gamma)$, with noncentrality parameter $\gamma \ge 0$ is defined as the distribution of the sum of squares $(Z_1+\gamma)^2 + Z_2 \cdots + Z_n^2$.

The random vector $\mathbf{Z}/R$ has length 1; it takes values on the unit sphere $S := \{\mathbf{z} \in \mathbb{R}^n : \sum_{i\le n} z_i^2 = 1\}$. By symmetry of the joint density $f(\mathbf{z})$, the random vector is uniformly distributed on S, no matter what value R takes. In other words $\mathbf{Z}/R$ is independent of R. This fact suggests a way to construct a random vector with the same distribution as $\mathbf{Z}$: Start with a random variable T^2 that has a χ_n^2 distribution independent of a random vector $\mathbf{U}$ that is uniformly distributed on the unit sphere S. Then the components of the random vector $T\mathbf{U}$ are independent $N(0,1)$'s. In two dimensions, the random vector $\mathbf{U}$ can be defined by

$$\mathbf{U} = (\cos V, \sin V) \qquad \text{where } V \sim \text{Unif}(0, 2\pi].$$

2.2 New coordinate system

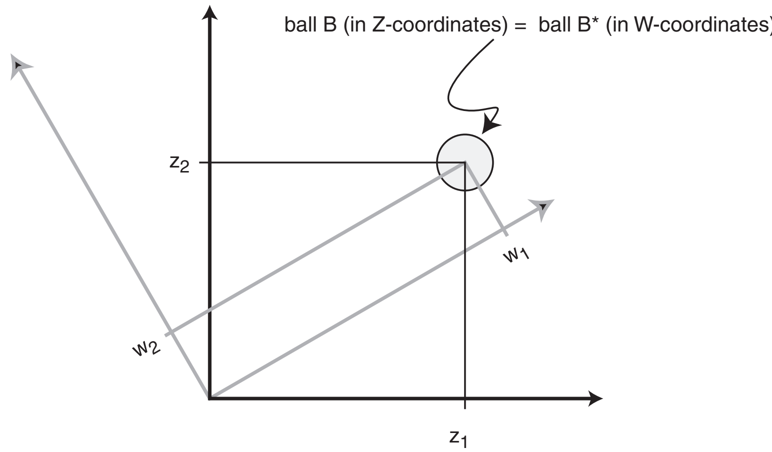
The spherical symmetry of the density $f(\cdot)$ is responsible for an important property of multivariate normals. Let $\mathbf{q}_1, \dots, \mathbf{q}_n$ be a new orthonormal basis for $\mathbb{R}^n$, and let

$$\mathbf{Z} = W_1\mathbf{q}_1 + \dots + W_n\mathbf{q}_n$$

be the representation for $\mathbf{Z}$ in the new basis.

<1> **Theorem.** *The $W_1, \dots, W_n$ are also independent $N(0,1)$ distributed random variables.*

If you know about multivariate characteristic functions this is easy to establish using the matrix representation $\mathbf{Z} = Q\mathbf{W}$, where Q is the orthogonal matrix with columns $\mathbf{q}_1, \dots, \mathbf{q}_n$.



A more intuitive explanation is based on the approximation

$$\mathbb{P}\{\mathbf{Z} \in B\} \approx f(\mathbf{z})(\text{volume of } B)$$

for a small ball B centered at $\mathbf{z}$. The transformation from $\mathbf{Z}$ to $\mathbf{W}$ corresponds to a rotation, so

$$\mathbb{P}\{\mathbf{Z} \in B\} = \mathbb{P}\{\mathbf{W} \in B^*\},$$

where B^* is a ball of the same radius, but centered at the point $\mathbf{w} = (w_1, \dots, w_n)$ for which $w_1\mathbf{q}_1 + \cdots + w_n\mathbf{q}_n = \mathbf{z}$. The last equality implies $\|\mathbf{w}\| = \|\mathbf{z}\|$, from which we get

$$\mathbb{P}\{\mathbf{W} \in B^*\} \approx (2\pi)^{-n/2} \exp(-\tfrac{1}{2}\|\mathbf{w}\|^2)(\text{volume of } B^*).$$

That is, $\mathbf{W}$ has the asserted spherical normal density.

To prove results about the spherical normal it is often merely a matter of transforming to an appropriate orthonormal basis.

<2> **Theorem.** *Let $\mathcal{X}$ be an m-dimensional subspace of $\mathbb{R}^n$. Let $\mathbf{Z}$ be a vector of independent $N(0,1)$ random variables, and $\boldsymbol{\mu}$ be a vector of constants. Then*

(i) the projection $\widehat{\mathbf{Z}}$ of $\mathbf{Z}$ onto $\mathcal{X}$ is independent of the projection $\mathbf{Z} - \widehat{\mathbf{Z}}$ of $\mathbf{Z}$ onto $\mathcal{X}^\perp$, the orthogonal complement of $\mathcal{X}$.

(ii) $\left\|\widehat{\mathbf{Z}}\right\|^2$ has a χ^2_m distribution.

(iii) $\|\mathbf{Z}+\boldsymbol{\mu}\|^2$ *has a noncentral* $\chi^2_n(\gamma)$ *distribution, with* $\gamma = \|\boldsymbol{\mu}\|$.

(iv) $\left\|\widehat{\mathbf{Z}}+\boldsymbol{\mu}\right\|^2$ *has a noncentral* $\chi^2_m(\gamma)$ *distribution, with* $\gamma = \|\boldsymbol{\mu}\|$.

PROOF Let $\mathbf{q}_1, \dots \mathbf{q}_n$ be an orthonormal basis of $\mathbb{R}^n$ such that $\mathbf{q}_1, \dots, \mathbf{q}_m$ span the space $\mathcal{X}$ and $\mathbf{q}_{m+1}, \dots, \mathbf{q}_n$ span $\mathcal{X}^\perp$. If $\mathbf{Z} = W_1\mathbf{q}_1 + \dots + W_n\mathbf{q}_n$ then

$$\begin{aligned} \widehat{\mathbf{Z}} &= W_1\mathbf{q}_1 + \dots + W_m\mathbf{q}_m, \\ \mathbf{Z} - \widehat{\mathbf{Z}} &= W_{m+1}\mathbf{q}_{m+1} + \dots + W_n\mathbf{q}_n, \\ \|\mathbf{Z}\|^2 &= W_1^2 + \dots + W_m^2, \end{aligned}$$

from which the first two asserted properties follow.

For the third and fourth assertions, choose the basis so that $\boldsymbol{\mu} = \gamma\mathbf{q}_1$. Then

$$\begin{aligned} \mathbf{Z} + \boldsymbol{\mu} &= (W_1+\gamma)\mathbf{q}_1 + W_2\mathbf{q}_2 + \dots \qquad\qquad + W_n\mathbf{q}_n \\ \widehat{\mathbf{Z}} + \boldsymbol{\mu} &= (W_1+\gamma)\mathbf{q}_1 + W_2\mathbf{q}_2 + \dots + W_m\mathbf{q}_m \end{aligned}$$

from which we get the noncentral chi-squares.

□

2.3 Fact about the general multivariate normal

If Z is an $n \times 1$ vector of independent $N(0,1)$ random variables, if μ is an $m \times 1$ vector of constants, and if A is an $m \times n$ matrix of constants, then the random vector $X = \mu + AZ$ has expected value μ and variance matrix $V = AA'$, and moment generating function

$$\mathbb{E}\exp(t'X) = \exp(t'\mu + t'AA't/2)$$

In particular, the distribution of X depends only on μ and V. The random vector X has a $N(\mu, V)$ distribution. If γ is a $k \times 1$ vector of constants and B is a $k \times m$ matrix of constants then

$$\gamma + BX = (\gamma + B\mu) + BAZ$$

has a $N(\gamma + B\mu, BVB')$ distribution.

2.4 Standard distributions

Suppose

Z has a $N(0,1)$ distribution
S_k^2 has a χ_k^2 distribution
S_ℓ^2 has a χ_ℓ^2 distribution

with all random variables independent of each other. Then, by definition,

$$\frac{Z}{\sqrt{S_k^2/k}} \text{ has a } t\text{-distribution on } k \text{ degrees of freedom } (t_k)$$

and

$$\frac{S_\ell^2/\ell}{S_k^2/k} \text{ has an } F\text{-distribution on } \ell \text{ and } k \text{ degrees of freedom } (F_{\ell,k})$$