Quadratic Forms

1.1 Definitions

Definition 1.1.1. For $\boldsymbol{x} \in \mathbb{R}^n$, a **quadratic form** with matrix $\boldsymbol{A}_{n\times n}$ is the function:

$$Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}.$$

Without loss of generality, $\boldsymbol{A}$ is assumed symmetric, since the quadratic form induced by a non-symmetric matrix $\tilde{\boldsymbol{A}}$ is equivalent to that induced by the symmetric matrix $\boldsymbol{A} = \frac{1}{2}(\tilde{\boldsymbol{A}} + \tilde{\boldsymbol{A}}')$. ■

Definition 1.1.2. A matrix $\boldsymbol{A}$ is **positive semi-definite** if for $\forall \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x} \neq \boldsymbol{0}$ the quadratic form $Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \geq 0$. If inequality is strict, then $\boldsymbol{A}$ is *positive definite.* ■

1.2 Properties

Remark 1.1.1. Recall from the spectral theorem that a symmetric matrix $\boldsymbol{A}_{n\times n}$ is uniquely expressible as $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'$, where $\boldsymbol{U}$ is an *orthogonal eigenbasis* and $\boldsymbol{\Lambda}$ is a diagonal matrix of *eigenvalues.* Since:

$$\boldsymbol{u}_i'\boldsymbol{u}_j = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

$\boldsymbol{A}$ may be viewed as a linear combination of projection operators:

$$\boldsymbol{A} = \sum_{i=1}^n \lambda_i \boldsymbol{u}_i\boldsymbol{u}_i' = \sum_{i=1}^n \lambda_i \boldsymbol{u}_i(\boldsymbol{u}_i'\boldsymbol{u}_i)^{-1}\boldsymbol{u}_i' = \sum_{i=1}^n \lambda_i \boldsymbol{P}_i,$$

where $\boldsymbol{P}_i = \boldsymbol{u}_i(\boldsymbol{u}_i'\boldsymbol{u}_i)^{-1}\boldsymbol{u}_i'$ is projection onto the image of the ith eigenvector $\boldsymbol{u}_i$. ◆

Proposition 1.1.1. For $\boldsymbol{x} \in \mathbb{R}^n$, the quadratic form $Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ is expressible as:

$$Q(\boldsymbol{x}) = \sum_{i=1}^n \lambda_i(\boldsymbol{u}_i'\boldsymbol{x})^2,$$

where (λ_i) are the eigenvalues and $(\boldsymbol{u}_i)$ the eigenvectors of $\boldsymbol{A}$. ◆

Proof. Let $\boldsymbol{A}_{n\times n} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'$. The quadratic form is expressible as:

$$Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = \text{tr}(\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}) = \text{tr}(\boldsymbol{x}'\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'\boldsymbol{x})$$

Let $\boldsymbol{y} = \boldsymbol{U}'\boldsymbol{x}$, then:

$$Q(\boldsymbol{x}) = \text{tr}(\boldsymbol{y}'\boldsymbol{\Lambda}\boldsymbol{y}) = \text{tr}(\boldsymbol{\Lambda}\boldsymbol{y}\boldsymbol{y}') = \sum_{i=1}^n \lambda_i y_i^2 = \sum_{i=1}^n \lambda_i(\boldsymbol{u}_i'\boldsymbol{x})^2$$

■

Proposition 1.1.2 (Rayleigh Quotient). For a symmetric matrix $\boldsymbol{A}_{n\times n}$:

$$\sup_{\{\boldsymbol{u}:||\boldsymbol{u}||^2=1\}} \boldsymbol{u}'\boldsymbol{A}\boldsymbol{u} = \max_{i=1}^{n}(\lambda_i),$$

where (λ_i) are the eigenvalues of $\boldsymbol{A}$. ♦

Proof. Let $Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ denote a quadratic form induced by $\boldsymbol{A}$. Take the spectral decomposition $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'$. Since $\boldsymbol{U}$ is an eigenbasis for $\mathbb{R}^n$, every $\boldsymbol{x}$ is expressible as $\boldsymbol{x} = \boldsymbol{U}\boldsymbol{\beta}_x$, for some coefficient $\boldsymbol{\beta}_x$. Now:

$$Q(\boldsymbol{x}) = \sum_{i=1}^{n}\lambda_i\big(\boldsymbol{u}_i'\boldsymbol{x}\big)^2 = \sum_{i=1}^{n}\lambda_i\big(\boldsymbol{u}_i'\boldsymbol{U}\boldsymbol{\beta}_x\big)^2 = \sum_{i=1}^{n}\lambda_i(\boldsymbol{e}_i'\boldsymbol{\beta}_x)^2 = \sum_{i=1}^{n}\lambda_i\beta_{x,i}^2 \leq \max_{i=1}^{n}(\lambda_i)\boldsymbol{\beta}_x'\boldsymbol{\beta}_x$$

The squared norm of the coefficient $||\boldsymbol{\beta}_x||^2$ is equal to the squared norm of the original vector $\boldsymbol{x}$:

$$||\boldsymbol{\beta}_x||^2 = \boldsymbol{\beta}_x'\boldsymbol{\beta}_x = \boldsymbol{\beta}_x'\boldsymbol{U}'\boldsymbol{U}\boldsymbol{\beta}_x = \big(\boldsymbol{U}\boldsymbol{\beta}_x\big)'\boldsymbol{U}\boldsymbol{\beta}_x = \boldsymbol{x}'\boldsymbol{x} = ||\boldsymbol{x}||^2.$$

Rearranging the bound on the quadratic form gives:

$$\frac{\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{\beta}_x'\boldsymbol{\beta}_x} = \frac{\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} \leq \max_{i=1}^{n}(\lambda_i)$$

Defining $\boldsymbol{u} = \boldsymbol{x}/||\boldsymbol{x}||$ such that $||\boldsymbol{u}||^2 = \boldsymbol{u}'\boldsymbol{u} = 1$ gives $\boldsymbol{u}'\boldsymbol{A}\boldsymbol{u} \leq \max(\lambda_i)$. Finally, since the bound is independent of $\boldsymbol{x}$, conclude that:

$$\sup_{\{\boldsymbol{u}:||\boldsymbol{u}||^2=1\}} \boldsymbol{u}'\boldsymbol{A}\boldsymbol{u} = \max_{i=1}^{n}(\lambda_i).$$

■

The χ^2_ν Distribution

2.1 Definitions

2.1.1 Central

Definition 1.2.1. If (Z_i) are independent $N(0,1)$ random variates, then the sum:

$$\sum_{i=1}^{\nu} Z_i^2 \stackrel{d}{=} \chi^2_\nu(0)$$

follows a central χ^2 distribution with ν degrees of freedom. ■

Proposition 1.2.1. The distribution of a $\chi^2_1(0)$ random variable is gamma with shape parameter $\alpha = 1/2$ and rate parameter $\lambda = 1/2$. ♦

Proof. Let $Z \sim N(0,1)$ and set $Y = Z^2$. The distribution of Y is:

$$P(Y \leq y) = P(Z^2 \leq y) = P(-\sqrt{y} \leq Z \leq \sqrt{y}) = \Phi(\sqrt{y}) - \Phi(-\sqrt{y}).$$

Differentiating to obtain the density:

$$f(y) = \phi(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} + \phi(-\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} = \phi(\sqrt{y}) \cdot \frac{1}{\sqrt{y}} = \frac{1}{\sqrt{2\pi}} y^{1/2-1} e^{-y/2}.$$

■

Proposition 1.2.2. The characteristic function of the $\chi^2_\nu(0)$ distribution is:

$$\phi(t) = (1 - 2i\omega)^{-\nu/2}.$$

♦

Proof. Suppose Z_i are IID $N(0,1)$. The characteristic function of $Y = \sum_{i=1}^{\nu} Z_i^2$ is:

$$\phi_\nu(\omega) = E\left(e^{i\omega \sum_{i=1}^n Z_i^2}\right) = \prod_{i=1}^{\nu} E\left(e^{i\omega Z_i^2}\right) = \prod_{i=1}^{\nu} \phi_1(\omega),$$

where $\phi_1(\omega)$ is the characteristic function of the $\chi^2_1(0)$ distribution. Finding $\phi_1(\omega)$:

$$\begin{aligned}\phi_1(\omega) = E\left(e^{i\omega Z_i^2}\right) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega z^2} e^{-z^2/2} dz \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}(1+2i\omega)} dz = \big(1 - 2i\omega\big)^{-1/2}.\end{aligned}$$

■

Corollary 1.2.1. The central χ^2 distribution with n degrees of freedom is a gamma distribution with shape parameter $\alpha = n/2$ and rate $\lambda = 1/2$.

$$\chi^2_n(0) \stackrel{d}{=} G(\alpha = n/2, \lambda = 1/2).$$

♣

2.1.2 Non-Central

Definition 1.2.2. If (X_i) are independent $N(\mu_i, \sigma^2)$ random variables, then the sum:

$$\frac{1}{\sigma^2} \sum_{i=1}^{\nu} X_i^2 \stackrel{d}{=} \chi^2_\nu(\delta)$$

follows a non-central χ^2 distribution with ν degrees of freedom and non-centrality parameter:

$$\delta = \frac{1}{\sigma^2} \sum_{i=1}^{\nu} \mu_i^2.$$

■

Proposition 1.2.3. For an integer $k \geq 0$:

$$2^{2k}(k!)\Gamma(k+1/2) = (2k)!\sqrt{\pi}.$$

♦

Proof. Using the recursion $\Gamma(k+1) = k\Gamma(k)$:

$$\begin{aligned}\Gamma(k+1/2) &= (k-1/2)(k-3/2)\cdots(1/2)\Gamma(1/2) \\ &= (1/2)^k(2k-1)(2k-3)\cdots(1)\sqrt{\pi} = (1/2)^k(2k-1)!!\sqrt{\pi}.\end{aligned}$$

Observe that $(2k-1)!!$ contains the odd terms from $(2k)!$, which is expressible as:

$$\begin{aligned}(2k)! &= 2k\cdot(2k-1)(2k-2)(2k-3)(2k-4)\cdots(2)(1) \\ &= (2k-1)(2k-3)\cdots(1)\cdot 2(k)2(k-1)\cdots 2(1) \\ &= (2k-1)!!\cdot 2^k(k!).\end{aligned}$$

Substituting $(2^k k!)^{-1}(2k)!$ for $(2k-1)!!$ in the expression for $\Gamma(k+1/2)$ gives:

$$\Gamma(k+1/2) = 2^{-k}(2^k k!)^{-1}(2k)!\sqrt{\pi}.$$

■

Proposition 1.2.4. The density of $\chi_1^2(\delta)$ is:

$$f(y) = e^{-\delta/2}\sum_{k=1}^{\infty}\frac{(\delta/2)^k}{k!}\cdot\frac{y^{k-1/2}e^{-y/2}}{2^{k+1/2}\Gamma(k+1/2)}.$$

♦

Proof. Suppose $Z \sim N(0,1)$, and let $Y = (Z+\mu)^2$, where $\mu = \sqrt{\delta}$. The distribution of Y is:

$$\begin{aligned}P(Y \leq y) &= P\big\{(Z+\mu)^2 \leq y\big\} = P(-\sqrt{y} \leq Z+\mu \leq \sqrt{y}) \\ &= P(-\sqrt{y}-\mu \leq Z \leq \sqrt{y}-\mu) = \Phi(\sqrt{y}-\mu) - \Phi(-\sqrt{y}-\mu).\end{aligned}$$

Differentiating to obtain the density:

$$\begin{aligned}f(y) &= \phi(\sqrt{y}-\mu)\cdot\frac{1}{2\sqrt{y}} + \phi(-\sqrt{y}-\mu)\cdot\frac{1}{2\sqrt{y}} \\ &= \frac{1}{2\sqrt{2\pi y}}\left(e^{-(\sqrt{y}-\mu)^2/2} + e^{-(\sqrt{y}+\mu)^2/2}\right) \\ &= \frac{e^{-(y+\mu^2)/2}}{2\sqrt{2\pi y}}\left(e^{\mu\sqrt{y}} + e^{-\mu\sqrt{y}}\right).\end{aligned}$$

Using the Taylor series representation:

$$\frac{1}{2}\left(e^{\mu\sqrt{y}} + e^{-\mu\sqrt{y}}\right) = \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{(\mu\sqrt{y})^k}{k!} + \frac{(-\mu\sqrt{y})^k}{k!}\right) = \frac{1}{2}\sum_{k=0}^{\infty}\frac{\sqrt{y}^k\mu^k}{k!}\{1+(-1)^k\}.$$

Observe that the summand vanishes for odd k, therefore:

$$\frac{1}{2}\sum_{k=0}^{\infty}\frac{\sqrt{y}^k\mu^k}{k!}\{1+(-1)^k\} = \sum_{l=0}^{\infty}\frac{\sqrt{y}^{2l}\mu^{2l}}{(2l)!} = \sum_{l=0}^{\infty}\frac{y^l\mu^{2l}}{(2l)!}.$$

Substituting the power series representation into the density:

$$f(y) = \frac{e^{-(y+\mu^2)/2}}{\sqrt{2\pi y}}\sum_{k=0}^{\infty}\frac{y^k\mu^{2k}}{(2k)!} = \frac{e^{-\mu^2/2}}{\sqrt{2}}\sum_{k=0}^{\infty}\frac{\mu^{2k}}{(2k)!\sqrt{\pi}}\cdot y^{k-1/2}e^{-y/2}.$$

Applying the identity from the previous proposition:

$$\begin{aligned} f(y) &= e^{-\mu^2/2}\sum_{k=0}^{\infty}\frac{\mu^{2k}}{2^{2k+1/2}(k!)\Gamma(k+1/2)}\cdot y^{k-1/2}e^{-y/2} \\ &= \sum_{k=0}^{\infty} e^{-\mu^2/2}\frac{(\mu^2/2)^k}{k!}\cdot\frac{1}{2^{k+1/2}\Gamma(k+1/2)}y^{k-1/2}e^{-y/2}. \end{aligned}$$

Observe that $e^{-\mu^2/2}(\mu^2/2)^k/(k!)$ is the density of the Poisson distribution, with mean $\mu^2/2$, and that:

$$\frac{1}{2^{k+1/2}\Gamma(k+1/2)}y^{k-1/2}e^{-y/2} = \frac{1}{2^{k+1/2}\Gamma(k+1/2)}y^{(2k+1)/2-1}e^{-y/2}$$

is the density of the $\chi^2_{2k+1}(0)$ distribution. The non-central $\chi^2_1(\delta)$ is therefore a Poisson mixture of central χ^2_{2k+1} distributions. ■

Proposition 1.2.5. The characteristic function of the $\chi^2_\nu(\delta)$ distribution is:

$$\phi(\omega) = e^{i\omega\delta/(1-2i\omega)}\cdot(1-2i\omega)^{-\nu/2}. \tag{1.2.1}$$

♦

Proof. Let (Z_i) denote independent $N(0,1)$, set $\mu = \sqrt{\delta}$, and define:

$$Y = (Z_\nu+\mu)^2 + \sum_{i=1}^{\nu-1} Z_i^2,$$

such that Y has a $\chi^2_\nu(\delta)$ distribution. The characteristic function of Y is:

$$\phi_Y(\omega) = E\left(e^{i\omega\left\{(Z_\nu+\mu)^2+\sum_{i=1}^{\nu-1}Z_i^2\right\}}\right) = E\left(e^{i\omega(Z_\nu+\mu)^2}\right)\cdot\prod_{i=1}^{\nu-1}E\left(e^{i\omega Z_i^2}\right) = \phi_{1,\delta}(\omega)\phi_1^{\nu-1}(\omega),$$

where $\phi_1(\omega)$ is the characteristic function of the $\chi_1^2(0)$ distribution, and $\phi_{1,\delta}(\omega)$ is the characteristic function of the $\chi_1^2(\delta)$ distribution. Finding the latter:

$$\phi_{1,\delta}(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\omega(z+\mu)^2}e^{-z^2/2}dz = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}g(z)}dz,$$

where:

$$\begin{aligned} g(z) &= z^2 - 2i\omega(z+\mu)^2 = z^2 - 2i\omega(z^2+\mu^2+2z\mu) \\ &= -2i\omega\mu^2 + z^2(1-2i\omega) - 4i\omega z\mu \\ &= -2i\omega\mu^2 + (1-2i\omega)\left\{z^2 - \frac{4i\omega z\mu}{1-2i\omega}\right\}. \end{aligned}$$

Let $\alpha = (1-2i\omega)^{-1}2i\omega\mu$, then:

$$\begin{aligned} g(z) &= -2i\omega\mu^2 + (1-2i\omega)\big\{z^2 - 2\alpha\big\} \\ &= -2i\omega\mu^2 + (1-2i\omega)\big\{(z-\alpha)^2 - \alpha^2\big\} \\ &= -2i\omega\mu^2 + \frac{4\omega^2\mu^2}{1-2i\omega} + (1-2i\omega)(z-\alpha)^2 \\ &= -2\omega\mu^2\left\{i - \frac{2\omega}{1-2i\omega}\right\} + (1-2i\omega)(z-\alpha)^2 \\ &= \frac{-2i\omega\mu^2}{1-2i\omega} + (1-2i\omega)(z-\alpha)^2. \end{aligned}$$

Thus, $\phi_{1,\delta}(\omega)$ resolves to:

$$\phi_{1,\delta}(\omega) = e^{i\omega\mu^2/(1-2i\omega)}\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-(1-2i\omega)(z-\alpha)^2/2}dz = e^{i\omega\mu^2/(1-2i\omega)}\cdot(1-2i\omega)^{-1/2}.$$

■

Corollary 1.2.2. The density of the $\chi_\nu^2(\delta)$ distribution is:

$$f(y) = e^{-\delta/2}\sum_{k=1}^{\infty}\frac{(\delta/2)^k}{k!}\cdot\frac{y^{(2k+\nu)/2-1}e^{-y/2}}{2^{k+\nu/2}\Gamma(k+\nu/2)}, \tag{1.2.2}$$

which is a Poisson $(\mu^2/2)$ mixture of $\chi_{2k+\nu}^2(0)$ distributions. ♣

Remark 1.2.1. The corollary is verified by showing the characteristic function of the $\chi_\nu^2(\delta)$ density in (1.2.2) is the characteristic function in (1.2.1). ♦

2.2 Properties

Proposition 1.2.6. The mean and variance of the $\chi_\nu^2(\delta)$ distribution are:

$$\begin{aligned} E\big\{\chi_\nu^2(\delta)\big\} &= \delta + \nu, \\ \mathrm{Var}\big\{\chi_\nu^2(\delta)\big\} &= 4\delta + 2\nu. \end{aligned}$$

♦

Proof. From the characteristic function, the moment generating function is:

$$M(t) = e^{t\delta/(1-2t)} \cdot (1-2t)^{-\nu/2}.$$

The cumulant generating function is:

$$K(t) = \frac{t\delta}{1-2t} - \frac{\nu}{2}\ln(1-2t).$$

The first derivative of K is:

$$\dot{K}(t) = \delta(1-2t)^{-1} + 2t\delta(1-2t)^{-2} + \nu(1-2t)^{-1}.$$

Evaluating at $t = 0$ gives the expectation:

$$E\big\{\chi^2_\nu(\delta)\big\} = \dot{K}(0) = \delta + \nu.$$

The second derivative of K is:

$$\ddot{K}(t) = 2\delta(1-2t)^{-2} + 2\delta(1-2t)^{-2} + 8t\delta(1-2t)^{-3} + 2\nu(1-2t)^{-2}.$$

Evaluating at $t = 0$ gives the variance:

$$\mathrm{Var}\big\{\chi^2_\nu(\delta)\big\} = \ddot{K}(0) = 4\delta + 2\nu.$$

■

Quadratic Forms in Normal Random Variables

Proposition 1.3.1. Suppose $\boldsymbol{y}_{n\times 1} \sim N\big(\boldsymbol{\mu}, \sigma^2\boldsymbol{I}_{n\times n}\big)$ and $\boldsymbol{P}_{n\times n}$ is the projection onto a linear subspace $V \subseteq \mathbb{R}^n$ of dimension $m = \dim(V) = \mathrm{tr}(\boldsymbol{P}) \leq n$. Then:

$$Q(\boldsymbol{y}) = \sigma^{-2} \cdot \boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} \sim \chi^2_m(\delta), \tag{1.3.3}$$

with non-centrality parameter (NCP) $\delta = \sigma^{-2} \cdot \boldsymbol{\mu}'\boldsymbol{P}\boldsymbol{\mu}$. ♦

Proof. Let $\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'$ denote the spectral decomposition of $\boldsymbol{P}$, then the quadratic form is expressible as: $Q(\boldsymbol{y}) = \sigma^{-2} \cdot \boldsymbol{y}'\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'\boldsymbol{y}$. Define $\boldsymbol{x} = \boldsymbol{U}'\boldsymbol{y}$, then since $\boldsymbol{U}$ is orthogonal, $\boldsymbol{x} \sim N\big(\boldsymbol{U}'\boldsymbol{\mu}, \sigma^2\boldsymbol{I}\big)$, which shows the components of $\boldsymbol{x}$ are independent.

Recall that all eigenvalues of a projection matrix are either 1 or 0. Since $\boldsymbol{P}$ is projection onto a subspace of dimension m, the first m eigenvalues are 1, and the remaining $(n-m)$ are 0. Now the quadratic form is expressible as:

$$Q(\boldsymbol{y}) = \sigma^{-2} \cdot \boldsymbol{x}'\boldsymbol{\Lambda}\boldsymbol{x} = \frac{1}{\sigma^2}\sum_{i=1}^{n}\lambda_i X_i^2 = \frac{1}{\sigma^2}\sum_{i=1}^{m} X_i^2.$$

where $X_i \sim N(\boldsymbol{u}_i'\boldsymbol{\mu}, \sigma^2)$. Conclude that $Q(\boldsymbol{y})$ follows a non-central $\chi_m^2(\delta)$ distribution, with non-centrality parameter:

$$\delta = \frac{1}{\sigma^2}\sum_{i=1}^{m}(\boldsymbol{u}_i'\boldsymbol{\mu})^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}\lambda_i(\boldsymbol{u}_i'\boldsymbol{\mu})^2 = \sigma^{-2}\cdot\boldsymbol{\mu}'\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'\boldsymbol{\mu} = \sigma^{-2}\cdot\boldsymbol{\mu}'\boldsymbol{P}\boldsymbol{\mu}.$$

■

Theorem 1.3.1. Suppose $\boldsymbol{y}_{n\times 1} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and $\boldsymbol{A}_{n\times n}$ is symmetric, positive definite. The distribution of $Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ is a mixture of non-central χ_1^2 distributions:

$$Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} \sim \sum_{i=1}^{n}\lambda_i\chi_1^2(\delta_i), \tag{1.3.4}$$

where $\delta_i = \nu_i^2$, $\nu_i = \boldsymbol{e}_i'\boldsymbol{U}'\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}$, $\boldsymbol{U}$ is the eigenbasis of $\boldsymbol{\Sigma}^{1/2}\boldsymbol{A}\boldsymbol{\Sigma}^{1/2}$, and (λ_i) are the corresponding eigenvalues. □

Proof. Let $\boldsymbol{y} = \boldsymbol{\Sigma}^{1/2}\boldsymbol{z} + \boldsymbol{\mu}$ where $\boldsymbol{z} \sim N(\boldsymbol{0}, \boldsymbol{I})$. Express $Q(\boldsymbol{y})$ as:

$$\begin{aligned}Q(\boldsymbol{y}) &= (\boldsymbol{\Sigma}^{1/2}\boldsymbol{z} + \boldsymbol{\mu})'\boldsymbol{A}(\boldsymbol{\Sigma}^{1/2}\boldsymbol{z} + \boldsymbol{\mu}) \\ &= (\boldsymbol{z} + \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu})'\boldsymbol{\Sigma}^{1/2}\boldsymbol{A}\boldsymbol{\Sigma}^{1/2}(\boldsymbol{z} + \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}).\end{aligned}$$

Let $\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'$ denote the spectral decomposition of $\boldsymbol{\Sigma}^{1/2}\boldsymbol{A}\boldsymbol{\Sigma}^{1/2}$. Re-expressing the quadratic form with use of the spectral decomposition of $\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'$:

$$\begin{aligned}Q(\boldsymbol{z}) &= (\boldsymbol{z} + \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu})'\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'(\boldsymbol{z} + \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}) \\ &= (\boldsymbol{U}'\boldsymbol{z} + \boldsymbol{U}'\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu})'\boldsymbol{\Lambda}(\boldsymbol{U}'\boldsymbol{z} + \boldsymbol{U}'\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}).\end{aligned}$$

Let $\boldsymbol{\nu} = \boldsymbol{U}'\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}$ and define $\boldsymbol{x} = \boldsymbol{U}'\boldsymbol{z} + \boldsymbol{\nu}$, then:

$$\boldsymbol{x} \sim \boldsymbol{U}'\boldsymbol{z} + \boldsymbol{\nu} \stackrel{d}{=} N\big(\boldsymbol{0}, \boldsymbol{I}\big) + \boldsymbol{\nu} \stackrel{d}{=} N\big(\boldsymbol{\nu}, \boldsymbol{I}\big).$$

It follows that the quadratic form is a weighted sum of non-central χ_1^2 distributions:

$$Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{\Lambda}\boldsymbol{x} = \sum_{i=1}^{n}\lambda_i X_i^2 = \sum_{i=1}^{n}\lambda_i\chi_1^2(\delta_i),$$

where $X_i \sim N\big(\nu_i, 1\big)$, $\nu_i = \boldsymbol{e}_i'\boldsymbol{\nu} = \boldsymbol{e}_i'\boldsymbol{U}'\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}$, and $\delta_i = \nu_i^2$. ■

Proposition 1.3.2. If $\boldsymbol{A}$ has a known Cholesky decomposition $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}'$, then the eigenbasis $\boldsymbol{U}$ and eigenvalues $\boldsymbol{\Lambda}$ appearing in (1.3.4) may be obtained from the spectral decomposition of $\boldsymbol{L}'\boldsymbol{\Sigma}\boldsymbol{L}$. ♦

Proof. Using the Cholesky decomposition:

$$Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{L}\boldsymbol{L}'\boldsymbol{y} = (\boldsymbol{L}'\boldsymbol{y})'\boldsymbol{I}(\boldsymbol{L}'\boldsymbol{y}).$$

Let $\boldsymbol{x} = \boldsymbol{L}'\boldsymbol{y} \sim N(\boldsymbol{L}'\boldsymbol{\mu}, \boldsymbol{L}'\boldsymbol{\Sigma}\boldsymbol{L})$. Applying (1.3.4), $\boldsymbol{U}$ is the eigenbasis of:

$$\big(\boldsymbol{L}'\boldsymbol{\Sigma}\boldsymbol{L}\big)^{1/2}\boldsymbol{I}\big(\boldsymbol{L}'\boldsymbol{\Sigma}\boldsymbol{L}\big)^{1/2} = \big(\boldsymbol{L}'\boldsymbol{\Sigma}\boldsymbol{L}\big).$$

■

Proposition 1.3.3. The matrix $\boldsymbol{\Sigma}^{1/2}\boldsymbol{A}\boldsymbol{\Sigma}^{1/2}$ appearing in (1.3.4) and the matrix $\boldsymbol{A}\boldsymbol{\Sigma}$ have the same eigenvalues. ♦

Proof. Recall that matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are *similar* if there exists an invertible *change of basis* $\boldsymbol{S}$ such that $\boldsymbol{B} = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$, and that similar matrices have the same eigenvalues. If $\boldsymbol{\Sigma}$ is positive definite, then $\boldsymbol{\Sigma}^{1/2}$ exists and is invertible; thus $\boldsymbol{A}\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}^{1/2}\boldsymbol{A}\boldsymbol{\Sigma}^{1/2}$ are similar matrices: $(\boldsymbol{A}\boldsymbol{\Sigma}) = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\Sigma}^{1/2}\boldsymbol{A}\boldsymbol{\Sigma}^{1/2})\boldsymbol{\Sigma}^{1/2}$. ■

Example 1.3.1. Suppose $\boldsymbol{\Sigma}_{n\times n}$ is symmetric, positive semi-definite, then by the spectral decomposition $\boldsymbol{\Sigma} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'$, where $\boldsymbol{U}$ is orthogonal and $\lambda_i \geq 0$ for each $i \in \{1, \cdots, n\}$. The square root of $\boldsymbol{\Sigma}$ is $\boldsymbol{\Sigma}^{1/2} = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}'$ since:

$$\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2} = \big(\boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}'\big)\big(\boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}'\big) = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}' = \boldsymbol{\Sigma}.$$

If $\boldsymbol{\Sigma}$ is positive definite ($\lambda_i > 0$ for $\forall i$), then the inverse of $\boldsymbol{\Sigma}$ is $\boldsymbol{\Sigma}^{-1} = \boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}'$ since:

$$\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1} = \big(\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'\big)\big(\boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}'\big) = \boldsymbol{I} = \big(\boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}'\big)\big(\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'\big) = \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}.$$

It follows that $\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}'$ is the square root of $\boldsymbol{\Sigma}^{-1}$. Moreover, expressing each matrix in terms of $\boldsymbol{U}$ and $\boldsymbol{\Lambda}$ gives the identities:

$$\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{I}, \qquad \boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^{1/2} = \boldsymbol{I}.$$

♠

Example 1.3.2. Suppose in (1.3.4) that $\boldsymbol{y}_{n\times 1} \sim N\big(\boldsymbol{\mu}, \boldsymbol{\Sigma}\big)$, and that $\boldsymbol{A} = \boldsymbol{\Sigma}^{-1}$. Then:

$$\boldsymbol{\Sigma}^{1/2}\boldsymbol{A}\boldsymbol{\Sigma}^{1/2} = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^{1/2} = \boldsymbol{I}.$$

Consequently, $\boldsymbol{U} = \boldsymbol{I}$ and $\lambda_i = 1$ for $\forall i$. Now $\nu_i = \boldsymbol{e}_i'\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}$, and:

$$Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} \sim \sum_{i=1}^{n}\chi_1^2(\delta_i) \stackrel{d}{=} \chi_n^2(\delta),$$
$$\delta = \sum_{i=1}^{n}\delta_i = \sum_{i=1}^{n}(\boldsymbol{e}_i'\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu})^2 = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}.$$

♠

Example 1.3.3. Suppose in (1.3.4) that $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}$, such that $\boldsymbol{y}_{n\times 1} \sim N\big(\boldsymbol{\mu}, \sigma^2\boldsymbol{I}\big)$. Moreover, let $\boldsymbol{x} = \sigma^{-1}\cdot\boldsymbol{y}$, such that $\boldsymbol{x}_{n\times 1} \sim N\big(\sigma^{-1}\boldsymbol{\mu}, \boldsymbol{I}\big)$. Consider the distribution of:

$$Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = \frac{\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}}{\sigma^2}.$$

By direct application of the theorem:

$$Q(\boldsymbol{x}) \sim \sum_{i=1}^{n}\lambda_i\chi_1^2(\delta_i),$$

where $\delta_i = \nu_i^2$, $\nu_i = \boldsymbol{e}_i'\boldsymbol{U}'\sigma^{-1}\boldsymbol{\mu} = \sigma^{-1}\boldsymbol{u}_i'\boldsymbol{\mu}$, $\boldsymbol{U}$ is the eigenbasis of $\boldsymbol{A}$, and (λ_i) are the corresponding eigenvalues. If in addition $\boldsymbol{A}$ is a projection matrix of dimension m, then the first m eigenvalues of $\boldsymbol{A}$ are 1, and the remainder 0, such that:

$$Q(\boldsymbol{x}) \sim \sum_{i=1}^{m}\chi_1^2(\delta_i) \overset{d}{=} \chi_m^2(\delta),$$

$$\delta = \sum_{i=1}^{m}\delta_i = \frac{1}{\sigma^2}\sum_{i=1}^{m}(\boldsymbol{u}_i'\boldsymbol{\mu})^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}\lambda_i\boldsymbol{\mu}'\boldsymbol{u}_i\boldsymbol{u}_i'\boldsymbol{\mu} = \frac{1}{\sigma^2}\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}.$$

Finally, if $\boldsymbol{\mu} = \boldsymbol{0}$ and $\boldsymbol{A} = \boldsymbol{I}$, which is a projection matrix, then:

$$Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{x} = \frac{\boldsymbol{y}'\boldsymbol{y}}{\sigma^2} \sim \sum_{i=1}^{n}\chi_1^2(0) \overset{d}{=} \chi_n^2(0).$$

♠

Remark 1.3.1. For numerical calculation of $P\big\{Q(\boldsymbol{x}) > t\big\}$, see the `CompQuadForm` package in `R`. Let $\boldsymbol{U}$ and $\boldsymbol{\Lambda}$ denote the eigenbasis and eigenvalues obtained from spectral decomposition of $\boldsymbol{\Sigma}^{1/2}\boldsymbol{A}\boldsymbol{\Sigma}^{1/2}$. The required inputs are the eigenvalues $\boldsymbol{\Lambda}$, and the non-centrality parameters $\boldsymbol{\delta} = \boldsymbol{\nu}^{\odot 2}$, where $\boldsymbol{\nu} = \boldsymbol{U}'\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}$ and $\boldsymbol{\nu}^{\odot 2}$ denotes the element-wise square. If $\boldsymbol{\mu} = \boldsymbol{0}$, then $\boldsymbol{\delta} = \boldsymbol{0}$, and the eigenvalues are more easily obtained from $\boldsymbol{A}\boldsymbol{\Sigma}$. ♦

Theorem 1.3.2 (Cochran's). Suppose $\boldsymbol{y}_{n\times 1} \sim N\big(\boldsymbol{\mu}, \boldsymbol{I}\big)$, and $(\boldsymbol{A}_i)$ are symmetric matrices with $\sum_i \boldsymbol{A}_i = \boldsymbol{I}$. Let $n_i = \text{rank}(\boldsymbol{A}_i)$ and $\delta_i = \boldsymbol{\mu}'\boldsymbol{A}_i\boldsymbol{\mu}$. Then, $Q_i = \boldsymbol{y}_i'\boldsymbol{A}_i\boldsymbol{y}_i$ are independent $\chi_{n_i}^2(\delta_i)$ random variables $\iff \sum_i n_i = n$. □

Example 1.3.4. Suppose $\boldsymbol{y}_{n\times 1} \sim N\big(\boldsymbol{\mu}, \sigma^2\boldsymbol{I}\big)$, and define $\boldsymbol{x}_{n\times 1} = \sigma^{-1}\boldsymbol{y} \sim N\big(\sigma^{-1}\boldsymbol{\mu}, \boldsymbol{I}\big)$. Let $\boldsymbol{P}$ denote projection onto a linear subpsace V of dimension m, and $\boldsymbol{Q} = \boldsymbol{I} - \boldsymbol{P}$ projection onto the orthogonal complement $V^{\perp}$. Now since $n_1 = m$ and $n_2 = n - m$, with $n_1 + n_2 = n$, by Cochran's theorem $\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} \sim \chi_m^2(\delta_1)$ and $\boldsymbol{y}'\boldsymbol{Q}\boldsymbol{y} \sim \chi_{n-m}^2(\delta_2)$ are independent, with $\delta_1 = \sigma^{-2}\cdot\boldsymbol{\mu}'\boldsymbol{P}\boldsymbol{\mu}$ and $\delta_2 = \sigma^{-2}\cdot\boldsymbol{\mu}'\boldsymbol{Q}\boldsymbol{\mu}$. ♠