2.3 Quadratic Forms $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$

2.3.1 Positive definiteness

A property of some quadratic forms used repeatedly in what follows is that of positive definiteness. A quadratic form is said to be *positive definite* if it is positive for all values of $\boldsymbol{x}$ except $\boldsymbol{x} = \mathbf{0}$; that is, if

$$\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} > 0 \quad \text{for all } \boldsymbol{x}, \text{ except } \boldsymbol{x} = \mathbf{0},$$

then $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} > 0$ is positive definite. And the corresponding (symmetric) matrix is also described as positive definite.

A slight relaxation of the above definition concerns $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ when its value is either positive or zero for all $\boldsymbol{x} \neq \mathbf{0}$. We define an $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ of this nature as being *positive semi-definite* (p.s.d) when

$$\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} > 0 \quad \text{for all } \boldsymbol{x} \neq \mathbf{0}, \text{ with } \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} = 0 \text{ for at least one } \boldsymbol{x} \neq \mathbf{0}.$$

Under these conditions, $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ is a p.s.d quadratic form and the corresponding symmetric matrix $\boldsymbol{A}$ is a p.s.d matrix. This definition is widely accepted; see Graybill (1976), Rao (1973), Scheffe (1959, p. 398).

The following are some properties of the positive (semi-) definite matrices, whose proofs can be found in the book "Linear models" by Searle and Gruber (Second edition).

1. The eigenvalues of a positive (semi-) definite matrix are all positive (non-negative). Thus, the determinant of a positive definite matrix is always positive, so a positive definite matrix is always nonsingular. A is p.d/p.s.d $\Rightarrow \lambda(A) >/\geq 0$
p.d $\Rightarrow |A| > 0$ A nonsingular.
2. A symmetric matrix is positive definite if and only if it can be written as $\boldsymbol{P}^\top \boldsymbol{P}$ for a non-singular matrix $\boldsymbol{P}$. A_i pd/psd $\Rightarrow \sum A_i$ p.d/psd
3. The sum of positive (semi) definite matrices is positive (semi-) definite.
4. A symmetric matrix $\boldsymbol{A}$ of order n and rank r can be written as $\boldsymbol{L}\boldsymbol{L}^\top$ when $\boldsymbol{L}$ is $n \times r$ of rank r, that is, $\boldsymbol{L}$ has full-column rank.

2.3.2 Quadratic forms

We discuss here the distribution of a quadratic form $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ when $\boldsymbol{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. We confine ourselves to the case where $\boldsymbol{A}$ is symmetric and $\boldsymbol{\Sigma}$ is non-singular. The following lemma presents the moment generating function of $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$.

LEMMA 1. *The M.G.F of $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ is* Recall noncentral χ^2 $\boxed{(1-2t)^{-\frac{n}{2}}}\boxed{e^{-\lambda[1-(1-2t)^{-1}]}}.$

$$M_{\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}}(t) = |I - 2tA\Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2}\boldsymbol{\mu}^\top[\boldsymbol{I}-(\boldsymbol{I}-2t\boldsymbol{A}\Sigma)^{-1}]\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}.$$

The proof of the above lemma can be found in textbook.

Recall Ch1 : X symmetric + $\lambda(X) \in \{0,1\} \Rightarrow X$ idempotent.

LEMMA 2. *If $\boldsymbol{A}$ and $\boldsymbol{V}$ are symmetric and $\boldsymbol{V}$ is positive definite, then $\boldsymbol{AV}$ has eigenvalues 0 and 1 implies that $\boldsymbol{AV}$ is idempotent.* $(AV)^\top = V^\top A^\top = VA$. not symmetric X.

Proof: Recall that that a symmetric matrix is positive definite if and only if it can be written as $\boldsymbol{P}^\top \boldsymbol{P}$ for a nonsingular $\boldsymbol{P}$. Then, $\boldsymbol{V} = \boldsymbol{P}^\top \boldsymbol{P}$ for some nonsingular matrix $\boldsymbol{P}$. If $|\boldsymbol{AV} - \lambda \boldsymbol{I}| = 0$ has roots 0 and 1, then

$$\begin{aligned} & |\boldsymbol{P}||\boldsymbol{AV} - \lambda\boldsymbol{I}||\boldsymbol{P}^{-1}| = 0 \qquad \text{has roots } 0 \,\&\, 1 \\ \Rightarrow\ & |\boldsymbol{PAVP^{-1}} - \lambda\boldsymbol{PIP^{-1}}| = 0 \qquad \text{has roots } 0 \,\&\, 1 \\ \Rightarrow\ & |\boldsymbol{PAP}^\top\boldsymbol{PP^{-1}} - \lambda\boldsymbol{I}| = 0 \qquad \text{has roots } 0 \,\&\, 1 \\ \Rightarrow\ & |\boldsymbol{PAP}^\top - \lambda\boldsymbol{I}| = 0 \qquad \text{has roots } 0 \,\&\, 1. \end{aligned} \tag{1}$$

Thus, $\boldsymbol{PAP}^\top$ has eigenvalues 0 and 1. But $\boldsymbol{PAP}^\top$ is symmetric (because $\boldsymbol{A}$ is symmetric). Hence, $\boldsymbol{PAP}^\top$ is idempotent. That is

$$\boldsymbol{PAP}^\top\boldsymbol{PAP}^\top = \boldsymbol{PAP}^\top \Rightarrow \boxed{\boldsymbol{PAVAP}^\top} = \boxed{\boldsymbol{PAP}^\top}.$$

As P is nonsingular,

$$\cancel{\boldsymbol{P}^{-1}}\boxed{\boldsymbol{PAVAP}^\top}\boldsymbol{P} = \cancel{\boldsymbol{P}^{-1}}\boxed{\boldsymbol{PAP}^\top}\boldsymbol{P}$$

which implies

$$\boldsymbol{AVAV} = \boldsymbol{AV}.$$

Consequently, we have proved $\boldsymbol{AV}$ is idempotent.

LEMMA 3. *If $\boldsymbol{A}$ is $n \times n$ symmetric idempotent matrix of rank r, then $\boldsymbol{A}$ has r eigenvalues equal to 1 and $n-r$ eigenvalues equal to 0.*

Next, we also review some properties of eigenvalues in order to prove the main theorems.

Properties of eigenvalues:

(1) If λ is an eigenvalue of $\boldsymbol{A}$, $c\lambda$ is an eigenvalue of $c\boldsymbol{A}$.

(2) For certain function $g(\boldsymbol{A})$, $g(\lambda)$ is an eigenvalue of $g(\boldsymbol{A})$.

(3) If $(\boldsymbol{I}-\boldsymbol{A})$ is nonsingular, then $1/(1-\lambda)$ is an eigenvalue of $(\boldsymbol{I}-\boldsymbol{A})^{-1}$.

(4) If $\underline{-1<\lambda<1,}$ then $1/(1-\lambda)$ can be represented by the series

$$\frac{1}{1-\lambda} = 1+\lambda+\lambda^2+\lambda^3+\cdots .$$

Correspondingly, if all eigenvalues of A satisfying $\underline{-1<\lambda<1,}$ then

$$(\boldsymbol{I}-\boldsymbol{A})^{-1} = \boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\boldsymbol{A}^3+\cdots .$$

(5) If $\boldsymbol{A}$ is any $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$, then $|\boldsymbol{A}| = \prod_{i=1}^n \lambda_i$ and $tr(\boldsymbol{A}) = \sum_{i=1}^n \lambda_i$.

dp & symm

THEOREM 1. *Let $\boldsymbol{x}_{p\times 1} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and let $\boldsymbol{A}$ be symmetric. Then, $q = \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ follows $\chi^2_{(r,\lambda)}$ with r being the rank of $\boldsymbol{A}$ and $\lambda = \frac{\boldsymbol{\mu}^\top \boldsymbol{A}\boldsymbol{\mu}}{2}$ <u>if and only of $\boldsymbol{A}\boldsymbol{\Sigma}$ is idempotent.</u>*

noncentral parameter

Proof: By Lemma 1, the moment generating function of $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ is

$$M_{\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}}(t) = \underbrace{\boxed{|\boldsymbol{I}-2t\boldsymbol{A}\boldsymbol{\Sigma}|}}_{}{}^{-\frac{1}{2}} e^{-\frac{1}{2}\mu^\top [\boxed{\boldsymbol{I}-(\boldsymbol{I}-2t\boldsymbol{A}\boldsymbol{\Sigma})^{-1}}]\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}$$

(1) (2)

with what conditions. it can equal to $(1-2t)^{-\frac{n}{2}} e^{\lambda[1-(1-2t)^{-1}]}$

Since the eigenvalue values of $\boldsymbol{I}-2t\boldsymbol{A}\boldsymbol{\Sigma}$ are $1-2t\lambda_i$, $i=1,\ldots,p$, where λ_i is an eigenvalue of $\boldsymbol{A}\boldsymbol{\Sigma}$. Then,

$$(1) \quad |\boldsymbol{I}-2t\boldsymbol{A}\boldsymbol{\Sigma}| = \prod_{i=1}^p (1-2t\lambda_i)$$

and

$$(2)\quad (\boldsymbol{I} - 2t\boldsymbol{A\Sigma})^{-1} = \boldsymbol{I} + \sum_{k=1}^{\infty}(2t\boldsymbol{A\Sigma})^k = I + \sum_{k=1}^{\infty}(2t)^k\boxed{(\boldsymbol{A\Sigma})^k}$$

if idempotent $\Rightarrow$ ...

idempotent $\Leftarrow$ ...

provided that $-1 < 2t\lambda_i < 1$ for all $i = 1, \ldots, p$. Hence,

$$\begin{aligned} M_{\boldsymbol{x}^\top \boldsymbol{Ax}}(t) &= \left\{\prod_{i=1}^{p}(1-2t\lambda_i)^{-\frac{1}{2}}\right\} exp\{-\frac{1}{2}\boldsymbol{\mu}^\top[-\sum_{k=1}^{\infty}(2t)^k]\boldsymbol{A\Sigma\Sigma^{-1}\mu}\} \\ &= (1-2t)^{-r/2}exp\{-\frac{1}{2}\boldsymbol{\mu}^\top\left(1-\frac{1}{1-2t}\right)\boldsymbol{A\mu}\} \end{aligned}$$

provided that $-1 < 2t < 1$ or $-1/2 < t < 1/2$ (which is compatible with the requirement that the M.G.F exists for t in a neighborhood of 0). This implies that

$$M_{\boldsymbol{x}^\top \boldsymbol{Ax}}(t) = \frac{1}{(1-2t)^{r/2}}exp\left\{-\frac{1}{2}\boldsymbol{\mu}^\top\boldsymbol{A\mu}\left(1-\frac{1}{1-2t}\right)\right\}$$

which is the M.G.F of a noncentral Chi-square random variable with degrees of freedom $r = rank(A)$ and noncentrality parameter $\lambda = \frac{1}{2}\boldsymbol{\mu}^\top\boldsymbol{A\mu}$.

For a proof of converse, if $\boldsymbol{x}^\top\boldsymbol{Ax}$ is $\chi^2(r,\lambda)$, then $\boldsymbol{A\Sigma}$ is idempotent. We leave space here for students' practice. Prove it inversely by myself □

The following are a few corollaries of Theorem 3.

Corollaries:

1. If $\boldsymbol{x} \sim N(\boldsymbol{0}, \boldsymbol{I})$, then $\boldsymbol{x}^\top\boldsymbol{Ax}$ is χ^2_r if and only if $\boldsymbol{A}$ is idempotent of rank r.

2. If $\boldsymbol{x} \sim N(\boldsymbol{0}, \boldsymbol{V})$, then $\boldsymbol{x}^\top\boldsymbol{Ax}$ is χ^2_r if and only if $\boldsymbol{AV}$ is idempotent of rank r.

3. If $\boldsymbol{x} \sim N(\boldsymbol{\mu}, \sigma^2\boldsymbol{I})$, then $\frac{\boldsymbol{x}^\top\boldsymbol{x}}{\sigma^2}$ is $\chi^2_{(n,\frac{1}{2\sigma^2}\boldsymbol{\mu}^\top\boldsymbol{\mu})}$.

4. If $\boldsymbol{x} \sim N(\boldsymbol{\mu}, \boldsymbol{I})$, the $\boldsymbol{x}^\top\boldsymbol{Ax}$ is $\chi^2_{(r,\frac{1}{2}\boldsymbol{\mu}^\top\boldsymbol{A\mu})}$ if and only if $\boldsymbol{A}$ is idempotent of rank r.

2.4 Independence

Let us first review more properties related to normal random vector. Suppose $\boldsymbol{x}_{p\times 1} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Properties of normal random vector:

(1) $E(\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}) = tr(\boldsymbol{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}^\top \boldsymbol{A}\boldsymbol{\mu}$;

(2) $Cov(\boldsymbol{x}, \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}) = 2\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\mu}$; ? $Cov(Bx, x^TAx) = 2B\Sigma A\mu$

(3) $Var(\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}) = 2tr[(\boldsymbol{A}\boldsymbol{V})^2] + 4\boldsymbol{\mu}^\top \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\mu}$. proof it by ourself.

The proofs of these properties are left as take-home exercises. $\square$

THEOREM 2. *Let $\boldsymbol{x} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Then $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ and $\boldsymbol{B}\boldsymbol{x}$ are distributed independently if and only if $\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{A} = \boldsymbol{0}$.*

$B\Sigma A \Rightarrow x^TAx \perp\!\!\!\perp Bx$

Proof: (*1. Sufficiency*) To prove $\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{A} = \boldsymbol{0}$ implies independence. Since $\boldsymbol{A}$ is symmetric, one can write $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^\top$ for some $\boldsymbol{L}$ of full column rank. If $\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{A} = \boldsymbol{0}$, then $\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{L}\boldsymbol{L}^\top = \boldsymbol{0}$. Next, as $\boldsymbol{L}$ has full column rank, $(\boldsymbol{L}^\top \boldsymbol{L})^{-1}$ exists. Thus,

$$\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{L}\boldsymbol{L}^\top = \boldsymbol{0} \Rightarrow \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{L}\boldsymbol{L}^\top \boldsymbol{L}(\boldsymbol{L}^\top \boldsymbol{L})^{-1} = \boldsymbol{0} \Rightarrow \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{L} = \boldsymbol{0}.$$

Therefore,

$$Cov(\boldsymbol{B}\boldsymbol{x}, \boldsymbol{x}^\top \boldsymbol{L}) = \boldsymbol{B}Cov(\boldsymbol{x})\boldsymbol{L} = \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{L} = \boldsymbol{0}.$$

Hence, since $\boldsymbol{x}$ is a vector of normally distributed random variables, $\boldsymbol{B}\boldsymbol{x}$ and $\boldsymbol{x}^\top \boldsymbol{L}$ are independently normally distributed which implies $\boldsymbol{B}\boldsymbol{x}$ and $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{L}^\top \boldsymbol{x}$ are distributed independently.

(*2. Necessity*) To prove independence of $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ and $\boldsymbol{B}\boldsymbol{x}$ implies $\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{A} = \boldsymbol{0}$. By independence, $Cov(\boldsymbol{B}\boldsymbol{x}, \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}) = \boldsymbol{0}$. On the other hand, $Cov(\boldsymbol{B}\boldsymbol{x}, \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}) = 2\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\mu}$. Hence,

$$2\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\mu} = \boldsymbol{0}.$$

it should be proofed for a given fixed μ not any μ.

Since this is true for all $\boldsymbol{\mu}$, $\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{A} = \boldsymbol{0}$ and the proof is complete.

Remark 1: In fact, the above proof of necessity appear in many textbooks is problematic and erroneous. "A correct proof for the general case of necessity requires the use of results from the theory of functions of complex variables (and depending on the approach taken, also from algebraic field theory). Consequently, the proof is inaccessible to many statisticians. Unfortunately, there is some evidence that a more accessible proof does not exist"; see Michael F. Driscoll, William R. Gundberg and Jr. (1986).

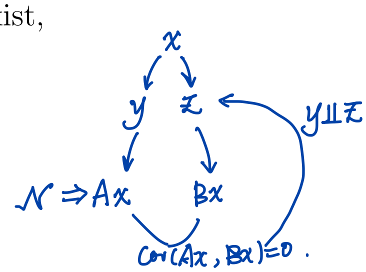
Remark 2: The above theorem does not involve $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B}$, a product that does not necessarily exist.

given μ . $B\Sigma A\mu = 0 \overset{?}{\Rightarrow} B\Sigma A = 0$.

THEOREM 3. *Let $\boldsymbol{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Then $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ and $\boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{x}$ are distributed independently if and only if $\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{A} = \boldsymbol{0}$(or equivalently, $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B} = \boldsymbol{0}$).*

Proof: (*1. Sufficiency*) That $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B} = \boldsymbol{0}$ implies independence. By Lemma 1, we can write $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^\top$ and $\boldsymbol{B} = \boldsymbol{M}\boldsymbol{M}^\top$ where each $\boldsymbol{L}$ and $\boldsymbol{M}$ have full column rank. Hence, if $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B} = \boldsymbol{0}$, then $\boldsymbol{L}\boldsymbol{L}^\top\boldsymbol{\Sigma}\boldsymbol{M}\boldsymbol{M}^\top = \boldsymbol{0}$. Since $(\boldsymbol{L}^\top\boldsymbol{L})^{-1}$ and $(\boldsymbol{M}^\top\boldsymbol{M})^{-1}$ exist,

$$(\boldsymbol{L}^\top\boldsymbol{L})^{-1}\boldsymbol{L}^\top(\boldsymbol{L}\boxed{\boldsymbol{L}^\top\boldsymbol{\Sigma}\boldsymbol{M}}\boldsymbol{M}^\top)\boldsymbol{M}(\boldsymbol{M}^\top\boldsymbol{M})^{-1} = \boldsymbol{0},$$



which implies

$$\boldsymbol{L}^\top\boldsymbol{\Sigma}\boldsymbol{M} = \boldsymbol{0}.$$

Therefore, $Cov(\boldsymbol{L}^\top\boldsymbol{x}, \boldsymbol{x}^\top\boldsymbol{M}) = \boldsymbol{L}^\top\boldsymbol{\Sigma}\boldsymbol{M} = \boldsymbol{0}$. Thus, as $\boldsymbol{x}$ is a vector of normally distributed variables, $\boldsymbol{L}^\top\boldsymbol{x}$ and $\boldsymbol{x}^\top\boldsymbol{M}$ are distributed independently. Consequently, $\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^\top\boldsymbol{L}\boldsymbol{L}^\top\boldsymbol{x}$ and $\boldsymbol{x}^\top\boldsymbol{B}\boldsymbol{x} = \boldsymbol{x}^\top\boldsymbol{M}\boldsymbol{M}^\top\boldsymbol{x}$ are distributed independently.

Similarly, the proof of necessity in many textbooks are erroneous. A relative complicate proof is available in the literature. We leave space here for students' exploration and future research for a simple proof.

THEOREM 4. *Let $\boldsymbol{x} \sim N(\boldsymbol{\mu}, \boldsymbol{V})$ and let $\boldsymbol{A}_i$ be $n \times n$ symmetric matrix of rank k_i, for $i = 1, 2, \cdots, p$. Denote $\boldsymbol{A} = \sum_{i=1}^{p} \boldsymbol{A}_i$, which is symmetric with rank k.*
Then,

$$\boldsymbol{x}^\top\boldsymbol{A_i}\boldsymbol{x} \sim \chi^2_{(k_i, \frac{1}{2}\boldsymbol{\mu}^\top\boldsymbol{A_i}\boldsymbol{\mu})},$$

$\boldsymbol{x}^\top\boldsymbol{A_i}\boldsymbol{x}$ are pairwise independent and $\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x}$ is $\chi^2_{(k, \frac{1}{2}\boldsymbol{\mu}^\top\boldsymbol{A}\boldsymbol{\mu})}$, if and only if

(I): any 2 of

(a) A_iV idempotent, for all i;

(b) $A_iVA_j = 0$ for all $i < j$;

(c) AV is idempotent.

Knowing how to apply them is enough for this course.

are true;

or (II): (c) is true and (d) $k = \sum_{i=1}^{p} k_i$;

or(III): (c) is true and (e) $A_1V, \cdots, A_{p-1}V$ are indempotent and A_pV is non-negative definite.

Corollary(Cochran's Theorem)

$x \sim N(0, \boldsymbol{I}_n)$ and A_i is symmetric of rank r_i for $i = 1, \cdots, p$ with $\sum_{i=1}^{p} A_i = \boldsymbol{I}_n$, then $x^\top A_i x$ are distributed independently as $\chi^2_{r_i}$ if and only if $\sum_{i=1}^{p} r_i = n$.

Some review of derivatives:

Let $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$ and let f be a function. $f : \mathbb{R}^p \to \mathbb{R}$

1. $\frac{\partial f}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_p} \end{bmatrix}$;
2. $\frac{\partial \boldsymbol{L}^\top x}{\partial \boldsymbol{x}} = \boldsymbol{L}$;
3. $\frac{\partial \boldsymbol{x}^\top \boldsymbol{L}}{\partial \boldsymbol{x}} = \boldsymbol{L}$;
4. $\frac{\partial \boldsymbol{x}^\top \boldsymbol{A} x}{\partial \boldsymbol{x}} = \boldsymbol{A}x + \boldsymbol{A}^\top \boldsymbol{x}$. (If $\boldsymbol{A}$ is symmetric, $\frac{\partial \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}}{\partial \boldsymbol{x}} = 2\boldsymbol{A}\boldsymbol{x}$.)