Quadratic forms

Let $\boldsymbol{A}$ be a real and symmetric $n \times n$ matrix. Then the *quadratic form* associated to $\boldsymbol{A}$ is the function $Q_{\boldsymbol{A}}$ defined by

$$Q_{\boldsymbol{A}}(\boldsymbol{x}) := \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \qquad (\boldsymbol{x} \in \mathbf{R}^n).$$

We have seen quadratic forms already, particularly in the context of positive-semidefinite matrices.

1. Random quadratic forms

Let $\boldsymbol{X} := (X_1, \ldots, X_n)'$ be an n-dimensional random vector. We are interested in the random quadratic form $Q_{\boldsymbol{A}}(\boldsymbol{X}) := \boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$.

Proposition 1. *If* $\mathrm{E}\boldsymbol{X} := \boldsymbol{\mu}$ *and* $\mathrm{Var}(\boldsymbol{X}) := \boldsymbol{\Sigma}$, *then*

$$\mathrm{E}\left(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\right) = \mathrm{tr}(\boldsymbol{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}.$$

In symbols, $\mathrm{E}(Q_{\boldsymbol{A}}(\boldsymbol{X})) = \mathrm{tr}(\boldsymbol{A}\boldsymbol{\Sigma}) + Q_{\boldsymbol{A}}(\boldsymbol{\mu})$.

Proof. We can write

$$\begin{aligned}\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X} &= (\boldsymbol{X}-\boldsymbol{\mu})'\boldsymbol{A}\boldsymbol{X} + \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{X}\\ &= (\boldsymbol{X}-\boldsymbol{\mu})'\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{\mu}) + \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{X} + (\boldsymbol{X}-\boldsymbol{\mu})'\boldsymbol{A}\boldsymbol{\mu}.\end{aligned}$$

If we take expectations, then the last term vanishes and we obtain

$$\mathrm{E}\left(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\right) = \mathrm{E}\left[(\boldsymbol{X}-\boldsymbol{\mu})'\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{\mu})\right] + \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}.$$

It suffices to verify that the expectation on the right-hand side is the trace of $\boldsymbol{A}\boldsymbol{\Sigma}$. But this is a direct calculation: Let $Y_j := X_i - \mu_j$, so that $\boldsymbol{Y} = \boldsymbol{X} - \boldsymbol{\mu}$

and hence

$$\begin{aligned}
\mathrm{E}\left[(\boldsymbol{X}-\boldsymbol{\mu})'\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{\mu})\right] &= \mathrm{E}\left(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}\right)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{E}\left(Y_iA_{i,j}Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}A_{i,j}\left[\mathrm{Var}(\boldsymbol{Y})\right]_{i,j}\\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}A_{i,j}\left[\mathrm{Var}(\boldsymbol{X}-\boldsymbol{\mu})\right]_{i,j} = \sum_{i=1}^{n}\sum_{j=1}^{n}A_{i,j}\left[\mathrm{Var}(\boldsymbol{X})\right]_{i,j}\\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}A_{i,j}\Sigma_{i,j} = \sum_{i=1}^{n}\sum_{i=1}^{n}A_{i,j}\Sigma_{j,i}\\
&= \sum_{i=1}^{n}\left[\boldsymbol{A}\boldsymbol{\Sigma}\right]_{i,i} = \mathrm{tr}(\boldsymbol{A}\boldsymbol{\Sigma}),
\end{aligned}$$

as desired. □

We easily get the following by a relabeling ($\boldsymbol{X} \Leftrightarrow \boldsymbol{X}-\boldsymbol{b}$):

Corollary 2. *For every nonrandom* $\boldsymbol{b}\in\mathbf{R}^n$*,*

$$\mathrm{E}\left[(\boldsymbol{X}-\boldsymbol{b})'\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{b})\right] = \mathrm{tr}(\boldsymbol{A}\boldsymbol{\Sigma}) + (\boldsymbol{\mu}-\boldsymbol{b})'\boldsymbol{A}(\boldsymbol{\mu}-\boldsymbol{b}).$$

In particular, $\mathrm{E}[(\boldsymbol{X}-\boldsymbol{\mu})'\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{\mu})] = \mathrm{tr}(\boldsymbol{A}\boldsymbol{\Sigma})$.

2. Examples of quadratic forms

What do quadratic forms look like? It is best to proceed by example.

Example 3. If $\boldsymbol{A} := \boldsymbol{I}_{n\times n}$, then $Q_{\boldsymbol{A}}(\boldsymbol{x}) = \sum_{i=1}^n x_i^2$. Because $\mathrm{tr}(\boldsymbol{A}\boldsymbol{\Sigma}) = \mathrm{tr}(\boldsymbol{\Sigma}) = \sum_{i=1}^n \mathrm{Var}(X_i)$, it follows that

$$\mathrm{E}\left(\sum_{i=1}^n X_i^2\right) = \sum_{i=1}^n \mathrm{Var}(X_i) + \left(\sum_{i=1}^n \mu_i^2\right).$$

This ought to be a familiar formula. □

Example 4. If

$$\boldsymbol{A} := \mathbf{1}_{m\times m} := \begin{pmatrix} 1 & \cdots & 1\\ \vdots & & \vdots\\ 1 & \cdots & 1\end{pmatrix}_{m\times m},$$

then $Q_{\boldsymbol{A}}(\boldsymbol{x}) = (\sum_{i=1}^n x_i)^2$. Note that

$$\mathrm{tr}(\boldsymbol{A}\boldsymbol{\Sigma}) = \sum_{i=1}^{n}\sum_{j=1}^{n}A_{i,j}\Sigma_{i,j} = \sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Cov}(X_i\,,X_j).$$

Therefore,

$$\mathrm{E}\left[\left(\sum_{i=1}^n X_i\right)^2\right] = \sum_{i=1}^n\sum_{j=1}^n \mathrm{Cov}(X_i\,, X_j) + \left(\sum_{i=1}^n \mu_i\right)^2;$$

this is another familiar formula. □

Example 5. One can combine matrices in a natural way to obtain new quadratic forms from old ones. Namely, if $a, b \in \mathbf{R}$ and $\boldsymbol{A}$ and $\boldsymbol{B}$ are real and symmetric $n \times n$ matrices, then $Q_{a\boldsymbol{A}+b\boldsymbol{B}}(\boldsymbol{x}) = aQ_{\boldsymbol{A}}(\boldsymbol{x}) + bQ_{\boldsymbol{B}}(\boldsymbol{x})$. For instance, suppose $\boldsymbol{A} := \boldsymbol{I}_{n\times n}$ and $\boldsymbol{B} := \boldsymbol{1}_{n\times n}$. Then,

$$a\boldsymbol{A} + b\boldsymbol{B} = \begin{pmatrix} a+b & b & b & \cdots & b \\ b & a+b & b & \cdots & b \\ b & b & a+b & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a+b \end{pmatrix},$$

and, thanks to the preceding two examples,

$$Q_{a\boldsymbol{A}+b\boldsymbol{B}}(\boldsymbol{x}) = a\sum_{i=1}^n x_i^2 + b\left(\sum_{i=1}^n x_i\right)^2.$$

An important special case is when $a := 1$ and $b := -1/n$. In that case,

$$\boldsymbol{A} - \frac{1}{n}\boldsymbol{B} = \begin{pmatrix} 1 - 1/n & -1/n & -1/n & \cdots & -1/n \\ -1/n & 1 - 1/n & -1/n & \cdots & -1/n \\ -1/n & -1/n & 1 - 1/n & \cdots & -1/n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1/n & -1/n & -1/n & \cdots & 1 - 1/n \end{pmatrix},$$

and

$$Q_{\boldsymbol{A}-(1/n)\boldsymbol{B}}(\boldsymbol{x}) = \sum_{i=1}^n x_i^2 - \frac{1}{n}\left(\sum_{i=1}^n x_i\right)^2 = \sum_{i=1}^n (x_i - \bar{x})^2.$$

Note that

$$\mathrm{tr}(\boldsymbol{A\Sigma}) = \sum_{i=1}^n \mathrm{Var}(X_i) + \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n \mathrm{Cov}(X_i\,, X_j).$$

Consider the special case that the X_i's are uncorrelated. In that case, $\mathrm{tr}(\boldsymbol{A\Sigma}) = (1 - 1/n)\sum_{i=1}^n \mathrm{Var}(X_i)$, and hence

$$\mathrm{E}\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = (1 - 1/n)\sum_{i=1}^n \mathrm{Var}(X_i) + \sum_{i=1}^n (\mu_i - \bar{\mu})^2.$$

When the X_i's are i.i.d. this yields $\mathrm{E}\sum_{i=1}^n (X_i - \bar{X})^2 = (n-1)\mathrm{Var}(X_1)$, which is a formula that you have seen in the context of the unbiasedness of the sample variance estimator $S^2 := (n-1)^{-1}\sum_{i=1}^n (X_i - \bar{X})^2$. □

Example 6. Consider a symmetric matrix of the form

$$\boldsymbol{A} := \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 1 & 0 & & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

That is, the super- and sub-diagonal entires are all ones and all other entries are zeros. Then,

$$Q_{\boldsymbol{A}}(\boldsymbol{x}) = 2\sum_{i=1}^{n-2} x_i x_{i+2}.$$

Other examples can be constructed in this way as well, and by also combining such examples. □

3. The variance of a random quadratic form

In the previous section we computed the expectation of $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ where $\boldsymbol{X}$ is a random vector. Here let us say a few things about the variance of the same random vector, under some conditions on $\boldsymbol{X}$.

Proposition 7. *Suppose $\boldsymbol{X} := (X_1, \ldots, X_n)'$ where the X_j's are i.i.d. with mean zero and four finite moments. Then,*

$$\mathrm{Var}\left(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\right) = \left(\mu_4 - 3\mu_2^2\right)\sum_{i=1}^n A_{i,i}^2 + \left(\mu_2^2 - 1\right)\left(\mathrm{tr}(\boldsymbol{A})\right)^2 + 2\mu_2^2 \mathrm{tr}\left(\boldsymbol{A}^2\right),$$

where $\mu_2 := \mathrm{E}(X_1^2)$ and $\mu_4 := \mathrm{E}(X_1^4)$.

One can generalize this a little more as well, with more or less the same set of techniques, in order to compute the variance of $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ in the case that the X_i's are independent, with common first four moments, and not necessarily mean zero.

Proof. Suppose $\boldsymbol{X} := (X_1, \ldots, X_n)'$, where $X_1, \ldots, X_n$ are independent and mean zero. Suppose $\mu_2 := \mathrm{E}(X_i^2)$ and $\mu_4 := \mathrm{E}(X_i^4)$ do not depend on i [e.g., because the X_j's are independent]. Then we can write

$$\left(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\right)^2 = \sum\sum\sum\sum_{1\le i,j,k,\ell\le n} A_{i,j}A_{k,\ell}X_iX_jX_kX_\ell.$$

Note that

$$\mathrm{E}\left(X_i X_j X_k X_\ell\right) = \begin{cases} \mu_4 & \text{of } i = j = k = \ell, \\ \mu_2^2 & \text{if } i = j \neq k = \ell \text{ or} \\ & \text{if } i = k \neq j = \ell \text{ or} \\ & \text{if } i = \ell \neq k = j, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore,

$$\begin{aligned} \mathrm{E}\left[\left(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\right)^2\right] &= \sum_{i=1}^n A_{i,i}^2\ \mu_4 + \sum\sum_{1\le i\neq k\le n} A_{i,i}A_{k,k}\ \mu_2^2 + \sum\sum_{1\le i\neq j\le n} A_{i,j}A_{j,i}\ \mu_2^2 + \sum\sum_{1\le i\neq k\le n} A_{i,k}A_{k,i}\ \mu_2^2 \\ &= \mu_4 \sum_{i=1}^n A_{i,i}^2 + \mu_2^2 \left[\sum\sum_{1\le i\neq k\le n} A_{i,i}A_{k,k} + 2\sum\sum_{1\le i\neq j\le n} A_{i,j}^2\right]. \end{aligned}$$

Next, we identify the double sums in turn:

$$\begin{aligned} \sum\sum_{1\le i\neq k\le n} A_{i,i}A_{k,k} &= \sum_{i=1}^n A_{i,i}\sum_{k=1}^n A_{k,k} - \sum_{i=1}^n A_{i,i}^2 = (\mathrm{tr}(\boldsymbol{A}))^2 - \sum_{i=1}^n A_{i,i}^2, \\ \sum\sum_{1\le i\neq j\le n} A_{i,j}^2 &= \sum_{i=1}^n\sum_{j=1}^n A_{i,j}^2 - \sum_{i=1}^n A_{i,i}^2 = \sum_{i=1}^n\sum_{j=1}^n A_{i,j}A_{j,i} - \sum_{i=1}^n A_{i,i}^2 \\ &= \sum_{i=1}^n (A^2)_{i,i} - \sum_{i=1}^n A_{i,i}^2 = \mathrm{tr}\left(\boldsymbol{A}^2\right) - \sum_{i=1}^n A_{i,i}^2. \end{aligned}$$

Consequently,

$$\begin{aligned} \mathrm{E}\left[\left(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\right)^2\right] &= \mu_4 \sum_{i=1}^n A_{i,i}^2 + \mu_2^2\left[(\mathrm{tr}(\boldsymbol{A}))^2 - \sum_{i=1}^n A_{i,i}^2 + 2\mathrm{tr}\left(\boldsymbol{A}^2\right) - 2\sum_{i=1}^n A_{i,i}^2\right] \\ &= \left(\mu_4 - 3\mu_2^2\right)\sum_{i=1}^n A_{i,i}^2 + \mu_2^2\left[(\mathrm{tr}(\boldsymbol{A}))^2 + 2\mathrm{tr}\left(\boldsymbol{A}^2\right)\right]. \end{aligned}$$

Therefore, in this case,

$$\mathrm{Var}\left(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\right) = \left(\mu_4 - 3\mu_2^2\right)\sum_{i=1}^n A_{i,i}^2 + \mu_2^2\left[(\mathrm{tr}(\boldsymbol{A}))^2 + 2\mathrm{tr}\left(\boldsymbol{A}^2\right)\right] - \left[\mathrm{E}\left(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\right)\right]^2.$$

This proves the result because $\mathrm{E}(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}) = \mathrm{tr}(\boldsymbol{A})$. □