4 Quadratic Forms I

Prof.Dr.P.M.Bajracharya

School of Mathematical Sciences
T.U., Kirtipur

April 22, 2024

Summary

1. Quadratic forms: Introduction
2. Reducing "Almost" Quadratic Forms to Quadratic Forms
3. Quadratic forms in terms of matrices
4. Some simple examples

Quadratic forms: Introduction

We have been viewing matrices in terms of **linear transformations**, but we can also view matrices in terms of **quadratic forms**.

Viewing matrices in terms of quadratic forms is a very powerful approach in machine learning and data science.

- It makes many advanced high-dimensional concepts more intuitive.
- It also provides a very nice geometric way to think about eigenvalues and eigenvectors that complements the more algebraic approach.

What does the term **quadratic** mean?
It is an expression of second degree.
For example a quadratic polynomial is a polynomial such as

$$ax^2 + bx + c.$$

An expression of the form

$$ax^2 + 2hxy + by^2,$$

where a, b and h are real numbers is called a **quadratic form in two variables** x and y.

Of course we can have a quadratic form of more than two variables as well.
We now place quadratic forms into matrix form as the examples below show.
Putting a quadratic into matrix form means we can use matrices to solve quadratic equations.

Example

Let $A = \begin{pmatrix} -2 & 4 \\ 4 & 3 \end{pmatrix}$ and $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Determine $x^T Ax$.

The matrix A is called the **matrix of quadratic form**.

Example

Let $A = \begin{pmatrix} a & h \\ h & b \end{pmatrix}$ and $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Determine $x^T A x$.

A quadratic form

A **quadratic form** $Q : \mathbb{R}^n \to \mathbb{R}$ is a polynomial in the variables $x_1, ..., x_n$, all of whose terms are of degree 2, that is,

$$Q(x) = x^T Ax = \sum_{i,j=1}^{n} Ax_i x_j,$$

where A is an $n \times n$ matrix.

Cross terms.

Terms of the form

$$x_i x_j \text{ , for } i \neq j$$

are sometimes called **cross terms**.

This is because they involve the product of two different variables, rather than the product of a variable with itself.

Diagonal terms.

Terms of the form x_i^2 , that involve the product of a variable with itself, are sometimes called **diagonal terms**.

A quadratic form can have diagonal terms or cross terms or both.

Diagonal quadratic forms.

Quadratic forms that do not have any cross terms are sometimes called **diagonal quadratic forms**.

Reducing "Almost" Quadratic to Quadratic Forms

An expression

$$4(x_1 - 3)^2 + x_2^2 = 4x_1^2 - 24x_1 + 36 - x_2^2$$

is not a quadratic form in x_1 and x_2.

An expression

$$4(x_1 - 3)^2 + x_2^2 = 4x_1^2 - 24x_1 + 36 - x_2^2$$

is not a quadratic form in x_1 and x_2. It is, however, "almost" a quadratic form.

Why "almost"?

An expression

$$4(x_1 - 3)^2 + x_2^2 = 4x_1^2 - 24x_1 + 36 - x_2^2$$

is not a quadratic form in x_1 and x_2. It is, however, "almost" a quadratic form.

Why "almost"?

The reason is that if we put $x_1' = x_1 - 3$, then

$$4(x_1 - 3)^2 + x_2^2 = 4x_1'^2 + x_2^2.$$

An expression

$$4(x_1 - 3)^2 + x_2^2 = 4x_1^2 - 24x_1 + 36 - x_2^2$$

is not a quadratic form in x_1 and x_2. It is, however, “almost” a quadratic form.

Why “almost”?

The reason is that if we put $x_1' = x_1 - 3$, then

$$4(x_1 - 3)^2 + x_2^2 = 4x_1'^2 + x_2^2.$$

By definition, the expression on the right is a quadratic in x_1' and x_2.

In general,

Redefining variables, i.e., performing a variable transformation, will permit us to remove those lower-order terms.

This will help us simplify/clarify the discussion.

"Removing" the linear and affine parts

Consider a function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = a + bx.$$

This function is the sum of
an afine part: a
and a linear part: bx.

We know that we can view this as a linear function by "removing" the afine part and considering the function

$$g(x) = f(x) - a = bx.$$

We know that we can view this as a linear function by "removing" the afine part and considering the function

$$g(x) = f(x) - a = bx.$$

This simply amounts to "shifting" or "translating" the function along the y-axis (if we view the function as $y = f(x)$), so that it goes through the origin.

Similarly, we can view the quadratic function given by the following equation

$$f(x) = cx^2 + bx + a$$

as a quadratic form by "removing" the linear and affine parts. To do so, one can use the procedure "completing the square".

We have

$$\begin{aligned} f(x) &= cx^2 + bx + a \\ &= c\left(x^2 + \frac{b}{c}x + \frac{a}{c}\right) \end{aligned}$$

We have

$$\begin{aligned} f(x) &= cx^2 + bx + a \\ &= c\left(x^2 + \frac{b}{c}x + \frac{a}{c}\right) \\ &= c\left(\left(x + \frac{b}{2c}\right)^2 - \frac{b^2 - 4ac}{4c^2}\right) \end{aligned}$$

We have

$$
\begin{aligned}
f(x) &= cx^2 + bx + a \\
&= c\left(x^2 + \frac{b}{c}x + \frac{a}{c}\right) \\
&= c\left(\left(x + \frac{b}{2c}\right)^2 - \frac{b^2 - 4ac}{4c^2}\right) \\
&= c\left(x + \frac{b}{2c}\right)^2 - \frac{b^2 - 4ac}{4c}
\end{aligned}
$$

We have

$$\begin{aligned} f(x) &= cx^2 + bx + a \\ &= c\left(x^2 + \frac{b}{c}x + \frac{a}{c}\right) \\ &= c\left(\left(x + \frac{b}{2c}\right)^2 - \frac{b^2 - 4ac}{4c^2}\right) \\ &= c\left(x + \frac{b}{2c}\right)^2 - \frac{b^2 - 4ac}{4c} \end{aligned}$$

and thus we have a quadratic with a vertex at

$$\left(-\frac{b}{2c}, -\frac{b^2 - 4ac}{4c}\right).$$

If we define

$$g(x) = f(x) + \frac{b^2 - 4ac}{4c}$$

$$x' = x + \frac{b}{2c},$$

then we have

$$g(x) = cx'^2,$$

which is a quadratic form in the variable x'.

Next, let's go to the two-variable case. Consider the function $f(x) : \mathbb{R}^2 \to \mathbb{R}$ given as

$$\begin{aligned} & f(x) = a + b_1x_1 + b_2x_2 \\ \Rightarrow \qquad & f(x) = a + b^T x, \end{aligned} \tag{1}$$

where $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$.

In this expression, b can be viewed in one of two complementary ways:

1. b is a **vector**, in which case $b^T x$ is a number that is the dot product of b and x.

In this expression, b can be viewed in one of two complementary ways:

1. b is a **vector**, in which case $b^T x$ is a number that is the dot product of b and x.
2. b^T is a 1×2 **matrix** that maps $x \in \mathbb{R}^2$ to a number in $\mathbb{R}$.

In this expression, b can be viewed in one of two complementary ways:

1. b is a **vector**, in which case $b^T x$ is a number that is the dot product of b and x.
2. b^T is a 1×2 **matrix** that maps $x \in \mathbb{R}^2$ to a number in $\mathbb{R}$.

In either case, the function f is the sum of an affine part (a) and a linear part ($b^T x$).

In this expression, b can be viewed in one of two complementary ways:

1. b is a **vector**, in which case $b^T x$ is a number that is the dot product of b and x.
2. b^T is a 1×2 **matrix** that maps $x \in \mathbb{R}^2$ to a number in $\mathbb{R}$.

In either case, the function f is the sum of an affine part (a) and a linear part ($b^T x$).
As before, we can convert this function using Equation (1) to a linear function by considering the translated function

$$g(x) = f(x) - a = b^T x.$$

Next, we consider a function $f: \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x) = a + b_1x_1 + b_2x_2 + c_1x_1^2 + c_2x_2^2 + c_3x_1x_2.$$

Next, we consider a function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x) = a + b_1x_1 + b_2x_2 + c_1x_1^2 + c_2x_2^2 + c_3x_1x_2.$$

Then we rewrite this equation as follows:

$$f(x) = a + (b_1 \ b_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (x_1 \ x_2)\begin{pmatrix} c_1 & c_3/2 \\ c_3/2 & c_2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Next, we consider a function $f: \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x) = a + b_1 x_1 + b_2 x_2 + c_1 x_1^2 + c_2 x_2^2 + c_3 x_1 x_2.$$

Then we rewrite this equation as follows:

$$\begin{aligned} f(x) &= a + (b_1 \ b_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (x_1 \ x_2)\begin{pmatrix} c_1 & c_3/2 \\ c_3/2 & c_2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= a + b^T x + x^T C x, \end{aligned} \tag{2}$$

where $C = \begin{pmatrix} c_1 & c_3/2 \\ c_3/2 & c_2 \end{pmatrix}$.

As before, we will be able to remove the linear and affine parts ($b^T x$ and a, respectively), again using the completing the square and shifting procedures. This too will lead to a quadratic form, but one with cross terms of the form $x_1 x_2$.

As before, we will be able to remove the linear and affine parts ($b^T x$ and a, respectively), again using the completing the square and shifting procedures. This too will lead to a quadratic form, but one with cross terms of the form $x_1 x_2$. Removing these cross terms in order to get a much simpler quadratic form without any cross terms will be closely related to computing *eigenvectors* and *eigenvalues*.

Generalization

To go to the three-variable case and beyond, observe that while Equation (1) and Equation (2) have been derived for a function $f : \mathbb{R}^2 \to \mathbb{R}$, exactly the same expressions could be derived for $f : \mathbb{R}^n \to \mathbb{R}$.

Generalization

To go to the three-variable case and beyond, observe that while Equation (1) and Equation (2) have been derived for a function $f : \mathbb{R}^2 \to \mathbb{R}$, exactly the same expressions could be derived for $f : \mathbb{R}^n \to \mathbb{R}$. The reason is that the equations just involve dot products and matrix-vector multiplications, and there is no explicit dependence on the dimension of the input to this function.

This suggests that we can write a general quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ in terms of an $n \times n$ symmetric matrix, a n-dimensional vector, and a number. This is true.

As an example, for $x \in \mathbb{R}^3$, the generalization of Equation (2) is

$$f(x) = a + (b_1\ b_2\ b_3)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + (x_1\ x_2\ x_3)\begin{pmatrix} c_{11} & c_{12}/2 & c_{13}/2 \\ c_{12}/2 & c_{22} & c_{23}/2 \\ c_{13}/2 & c_{23}/2 & c_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{3}$$

Quadratic forms in terms of matrices

The above discussion gives rise of two problems:

Problem I. Given a matrix, find the associated quadratic form.

Problem II. Given a quadratic form, find the associated matrix.

Problem I

Let's start with an arbitrary square matrix. Starting with an arbitrary 3×3 matrix for simplicity,

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \in \mathbb{R}^{3\times 3},$$

where, in particular, we permit the possibility that $A_{12} = A_{21}$ as well as the possibility that $A_{12} \neq A_{21}$, and similarly for the other off-diagonal terms.

In this case, we have

$$Ax = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

In this case, we have

$$
\begin{aligned}
Ax &= \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
&= \begin{pmatrix} A_{11}x_1 + A_{12}x_2 + A_{13}x_3 \\ A_{21}x_1 + A_{22}x_2 + A_{23}x_3 \\ A_{31}x_1 + A_{32}x_2 + A_{33}x_3 \end{pmatrix} \in \mathbb{R}^3.
\end{aligned}
$$

And we also have

$$x^T A x = (x_1 \; x_2 \; x_3) \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

And we also have

$$x^T A x = (x_1 \ x_2 \ x_3) \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= A_{11}x_1^2 + (A_{12} + A_{21})x_1x_2 + (A_{13} + A_{31})x_1x_3$$
$$+ A_{22}x_2^2 + (A_{23} + A_{32})x_2x_3 + A_{33}x_3^2 \in \mathbb{R}^3.$$

And we also have

$$x^T A x = (x_1\ x_2\ x_3)\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$
$$= A_{11}x_1^2 + (A_{12} + A_{21})x_1x_2 + (A_{13} + A_{31})x_1x_3$$
$$+ A_{22}x_2^2 + (A_{23} + A_{32})x_2x_3 + A_{33}x_3^2 \in \mathbb{R}^3.$$

The point of this is to show that we can start with an arbitrary square matrix and get a quadratic form (i.e., we don't need to impose any condition on a matrix to have a quadratic form) .

Problem II

Let's go the "other way," i.e., let's start with an arbitrary quadratic form.

Problem II

Let's go the "other way," i.e., let's start with an arbitrary quadratic form.

Let's consider the case where the function $f : \mathbb{R}^3 \to \mathbb{R}$ has just quadratic terms, i.e., where it is a quadratic form.

Problem II

Let's go the "other way," i.e., let's start with an arbitrary quadratic form.

Let's consider the case where the function $f : \mathbb{R}^3 \to \mathbb{R}$ has just quadratic terms, i.e., where it is a quadratic form. In this case, we are considering

$$\begin{aligned} f(x) = c_{11}x_1^2 + c_{22}x_2^2 + c_{33}x_3^2 \\ + c_{12}x_1x_2 + c_{13}x_1x_3 + c_{23}x_2x_3. \end{aligned}$$

We have

$$\begin{aligned} f(x) = x_1(&c_{11}x_1 + c_{12}x_2 + c_{13}x_3) \\ &+ x_2(c_{22}x_2 + c_{23}x_3) + c_{33}x_3^2 \end{aligned}$$

We have

$$
\begin{aligned}
f(x) &= x_1(c_{11}x_1 + c_{12}x_2 + c_{13}x_3) \\
&\qquad + x_2(c_{22}x_2 + c_{23}x_3) + c_{33}x_3^2 \\
&= (x_1 \ x_2 \ x_3) \begin{pmatrix} c_{11}x_1 + c_{12}x_2 + c_{13}x_3 \\ c_{22}x_2 + c_{23}x_3 \\ c_{33} \end{pmatrix}
\end{aligned}
$$

We have

$$
\begin{aligned}
f(x) &= x_1(c_{11}x_1 + c_{12}x_2 + c_{13}x_3) \\
&\quad + x_2(c_{22}x_2 + c_{23}x_3) + c_{33}x_3^2 \\
&= (x_1\ x_2\ x_3)\begin{pmatrix} c_{11}x_1 + c_{12}x_2 + c_{13}x_3 \\ c_{22}x_2 + c_{23}x_3 \\ c_{33} \end{pmatrix} \\
&= (x_1\ x_2\ x_3)\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.
\end{aligned}
$$

Therefore,

$$f(x) = (x_1 \ x_2 \ x_3) \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x^T A x.$$

Therefore,

$$f(x) = (x_1\ x_2\ x_3)\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x^T A x.$$

It follows that

$$A = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{pmatrix}$$

is the matrix associated with the given quadratic form.

Thus, the point here is that we can start with an arbitrary quadratic form and construct in a very natural way a matrix associated with the given quadratic form.

Thus, the point here is that we can start with an arbitrary quadratic form and construct in a very natural way a matrix associated with the given quadratic form.

Non-uniqueness

However, even if we restrict ourselves to quadratic functions that contain just terms of degree 2, this quadratic form does not uniquely define a matrix.

For example, we could have put the c_{12}, c_{13} and c_{23} below the diagonal, with 0s above the diagonal as follows:

$$f(x) = (x_1\ x_2\ x_3)\begin{pmatrix} c_{11} & 0 & 0 \\ c_{12} & c_{22} & 0 \\ c_{13} & c_{23} & c_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x^T A^T x.$$

For example, we could have put the c_{12}, c_{13} and c_{23} below the diagonal, with 0s above the diagonal as follows:

$$f(x) = (x_1\, x_2\, x_3) \begin{pmatrix} c_{11} & 0 & 0 \\ c_{12} & c_{22} & 0 \\ c_{13} & c_{23} & c_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x^T A^T x.$$

It follows that

$$A^T = \begin{pmatrix} c_{11} & 0 & 0 \\ c_{12} & c_{22} & 0 \\ c_{13} & c_{23} & c_{33} \end{pmatrix}$$

is the matrix associated with the given quadratic form.

Alternatively, we can write it in a more symmetric form as follows:

$$f(x) = (x_1\ x_2\ x_3)\begin{pmatrix} c_{11} & c_{12}/2 & c_{13}/2 \\ c_{12}/2 & c_{22} & c_{23}/2 \\ c_{13}/2 & c_{23}/2 & c_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Alternatively, we can write it in a more symmetric form as follows:

$$f(x) = (x_1\ x_2\ x_3)\begin{pmatrix} c_{11} & c_{12}/2 & c_{13}/2 \\ c_{12}/2 & c_{22} & c_{23}/2 \\ c_{13}/2 & c_{23}/2 & c_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

In this last expression, we have split the "off diagonal" terms, i.e., c_{ij} , for $i \neq j$, into half associated with the $x_i x_j$ term and half associated with the $x_j x_i$ term.

Right now, we just want to note that this is possible, i.e., we didn't have to do this for the above expression to be true (as we saw with the previous two non-symmetric examples that also reproduce the same $f(x)$).

Right now, we just want to note that this is possible, i.e., we didn't have to do this for the above expression to be true (as we saw with the previous two non-symmetric examples that also reproduce the same $f(x)$). It turns out, however, to be very convenient to do this. We'll get back to why this is the case soon.

Clearly, if we define the matrix A' to be

$$A' = \begin{pmatrix} A_{11} & \frac{A_{12}+A_{21}}{2} & \frac{A_{13}+A_{31}}{2} \\ \frac{A_{12}+A_{21}}{2} & A_{22} & \frac{A_{23}+A_{32}}{2} \\ \frac{A_{13}+A_{31}}{2} & \frac{A_{23}+A_{32}}{2} & A_{33} \end{pmatrix} \in \mathbb{R}^{3\times 3},$$

then $x^T Ax = x^T A'x$, for all $x \in \mathbb{R}^3$, illustrating again the same non-uniqueness.
Here is a specific example of two matrices that correspond to the same quadratic form.

Example. Let

$$A_1 = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 6 & 8 \\ 0 & 0 & 9 \end{pmatrix} \text{ and } A_2 = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 6 & 4 \\ 2 & 4 & 9 \end{pmatrix}.$$

Then

$$x^T A_1 x = (x_1 \; x_2 \; x_3) \begin{pmatrix} 1 & 2 & 4 \\ 0 & 6 & 8 \\ 0 & 0 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$
$$= x_1^2 + 6x_2^2 + 9x_3^2 + 2x_1x_2 + 4x_1x_3 + 8x_2x_3$$

Then

$$
\begin{aligned}
x^T A_1 x &= (x_1\ x_2\ x_3) \begin{pmatrix} 1 & 2 & 4 \\ 0 & 6 & 8 \\ 0 & 0 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
&= x_1^2 + 6x_2^2 + 9x_3^2 + 2x_1x_2 + 4x_1x_3 + 8x_2x_3 \\
&= x_1^2 + 6x_2^2 + 9x_3^2 + (x_1x_2 + x_2x_1) \\
&\qquad + (2x_1x_3 + 2x_3x_1) + (4x_2x_3 + 4x_3x_2)
\end{aligned}
$$

Then

$$
\begin{aligned}
x^T A_1 x &= (x_1 \ x_2 \ x_3) \begin{pmatrix} 1 & 2 & 4 \\ 0 & 6 & 8 \\ 0 & 0 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
&= x_1^2 + 6x_2^2 + 9x_3^2 + 2x_1x_2 + 4x_1x_3 + 8x_2x_3 \\
&= x_1^2 + 6x_2^2 + 9x_3^2 + (x_1x_2 + x_2x_1) \\
&\qquad + (2x_1x_3 + 2x_3x_1) + (4x_2x_3 + 4x_3x_2) \\
&= (x_1 \ x_2 \ x_3) \begin{pmatrix} 1 & 1 & 2 \\ 1 & 6 & 4 \\ 2 & 4 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
&= x^T A_2 x.
\end{aligned}
$$

This discussion illustrates that we can have many different square matrices that give rise to the same quadratic form.

This discussion illustrates that we can have many different square matrices that give rise to the same quadratic form.

We might wonder whether there is some sort of standard or canonical form in which we can write a matrix that removes this non-uniqueness.

This discussion illustrates that we can have many different square matrices that give rise to the same quadratic form.

We might wonder whether there is some sort of standard or canonical form in which we can write a matrix that removes this non-uniqueness.

If it is possible, then we can write any quadratic form in terms of a unique matrix.

The answer to this is

yes.

The answer to this is

yes.

Basically,

The way to do this is to do the above "average out" and write the quadratic form as a symmetric matrix (like we did above).

The answer to this is

yes.

Basically,

The way to do this is to do the above "average out" and write the quadratic form as a symmetric matrix (like we did above).

That is, consider **symmetric matrices**. We could do it other ways, e.g., put everything above the diagonal, but we will do it this way since symmetric matrices have so many nice properties.

Conversely,

For any symmetric matrix, there is an associated quadratic form.

The point is the following.

In general,

If A is a square $n \times n$ matrix, then the function $f : \mathbb{R}^n \to \mathbb{R}$ given by

$$\boldsymbol{f(x) = x^T A x}$$

is a quadratic form.

The point is the following.

In general,

If A is a square $n \times n$ matrix, then the function $f : \mathbb{R}^n \to \mathbb{R}$ given by

$$\boldsymbol{f(x) = x^T A x}$$

is a quadratic form.

In other words, a quadratic form is a product of three matrices, two of which are vectors, that for a given x and A yields a 1×1 matrix that is a number.

Theorem.

If $A \in \mathbb{R}^{3\times 3}$ with a quadratic form in 3 variables, then there is a symmetric matrix $B \in \mathbb{R}^{3\times 3}$ such that

$$\forall\, x \in \mathbb{R}^3 \quad x^T A x = x^T B x.$$

Proof.

Let

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}.$$

Proof...

Then the quadratic form is given by

$$x^T Ax = (x_1\ x_2\ x_3)\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Proof...

Then the quadratic form is given by

$$x^T Ax = (x_1 \ x_2 \ x_3) \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= A_{11}x_1^2 + (A_{12} + A_{21})x_1x_2 + (A_{13} + A_{31})x_1x_3$$

Proof...

Then the quadratic form is given by

$$x^T A x = (x_1\ x_2\ x_3) \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= A_{11}x_1^2 + (A_{12} + A_{21})x_1x_2 + (A_{13} + A_{31})x_1x_3$$

$$+ A_{22}x_2^2 + (A_{23} + A_{32})x_2x_3 + A_{33}x_3^2 \in \mathbb{R}^3.$$

Proof...

We put

$$B_{ij} = \frac{A_{ij} + A_{ji}}{2}.$$

Proof...

We put

$$B_{ij} = \frac{A_{ij} + A_{ji}}{2}.$$

Then the matrix $B = (B_{ij})_{3\times 3}$ is symmetric, since

$$B_{ij} = \frac{A_{ij} + A_{ji}}{2} = \frac{A_{ji} + A_{ij}}{2} = B_{ji}.$$

Proof...

Moreover,

$$B_{ij} + B_{ji} = \frac{A_{ij} + A_{ji}}{2} + \frac{A_{ji} + A_{ij}}{2} = A_{ij} + A_{ji}.$$

Proof...

Moreover,

$$B_{ij} + B_{ji} = \frac{A_{ij} + A_{ji}}{2} + \frac{A_{ji} + A_{ij}}{2} = A_{ij} + A_{ji}.$$

Therefore,

$$\forall\ x \in \mathbb{R}^3 \quad x^T A x = x^T B x.$$

$\square$

In general we have the following theorem.

Theorem.

Given an arbitrary quadratic form in n variables (which, recall can be written as $x^T Ax$ for a square matrix $A \in \mathbb{R}^{n\times n}$), we can always find a symmetric matrix $B \in \mathbb{R}^{n\times n}$) such that

$$\forall\ x \in \mathbb{R}^n \quad x^T Ax = x^T Bx.$$

:

In general we have the following theorem.

Theorem.

Given an arbitrary quadratic form in n variables (which, recall can be written as $x^T Ax$ for a square matrix $A \in \mathbb{R}^{n\times n}$), we can always find a symmetric matrix $B \in \mathbb{R}^{n\times n}$) such that

$$\forall\, x \in \mathbb{R}^n \quad x^T Ax = x^T Bx.$$

:

Due to this result, nothing is lost if we assume that A is symmetric. So, **from now on, we do this**.

By the way, we agree always to choose the symmetric matrix for two reasons.

1. It gives us a unique (symmetric) matrix corresponding to a given quadratic form.
2. The choice of the symmetric matrix A allows us to apply the special theory available for symmetric matrices.

Note that

From any matrix, we get a type of quadratic form using
$$x^T Ay.$$
In this case, the variables x and y have different dimensions.

For example, if

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix},$$

then we obtain the following:

$$x^T A x = (x_1 \ x_2) \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

For example, if

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix},$$

then we obtain the following:

$$x^T A x = (x_1 \ x_2) \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$= A_{11}x_1y_1 + A_{12}x_1y_2 + A_{13}x_1y_3$$

For example, if

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix},$$

then we obtain the following:

$$\begin{aligned} x^T A x &= (x_1 \; x_2) \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ &= A_{11}x_1y_1 + A_{12}x_1y_2 + A_{13}x_1y_3 \\ &\qquad + A_{21}x_2y_1 + A_{22}x_2y_2 + A_{23}x_2y_3 \in \mathbb{R}^3. \end{aligned}$$

Instead, we will focus on square matrices, in which case the vectors on the two sides of the matrix can be represented by the same variable.

Some simple examples

Problem. Find the quadratic form corresponding to each of the following symmetric matrices.

1. $\begin{pmatrix} 4 & 1/\sqrt{2} \\ 1/2 & \sqrt{2} \end{pmatrix}$
2. $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 3 \\ 0 & 3 & -1 \end{pmatrix}$

Problem. Find the corresponding symmetric matrix for each of the following quadratic forms.

1. $Q(x) = x_1^2 - 2x_1x_2 - 3x_2^2$
2. $Q(x) = 2x_1^2 + 3x_1x_2 - x_1x_3 + 4x_2^2 + x_3^2$
3. $Q(x) = x_1^2 + 2x_2^2 + 3x_3^2 + 4x_4^2$

Example 1.

Let $Q(x) = 8x_1^2 - 4x_1x_2 + 5x_2^2$. Determine whether $Q(0,0)$ is the global minimum.

Solution. We can rewrite the given equation in quadratic form as follows:

$$Q(x) = x^T Ax, \text{ where } A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}.$$

Solution. We can rewrite the given equation in quadratic form as follows:

$$Q(x) = x^T Ax, \text{ where } A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}.$$

Then

$$\begin{vmatrix} 8-\lambda & -2 \\ -2 & 5-\lambda \end{vmatrix} = 0$$

Solution. We can rewrite the given equation in quadratic form as follows:

$$Q(x) = x^T A x, \text{ where } A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}.$$

Then

$$\begin{vmatrix} 8-\lambda & -2 \\ -2 & 5-\lambda \end{vmatrix} = 0$$

$$\Rightarrow \quad (8-\lambda)(5-\lambda) - 4 = 0$$

Solution. We can rewrite the given equation in quadratic form as follows:

$$Q(x) = x^T Ax, \text{ where } A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}.$$

Then

$$\begin{aligned} & \begin{vmatrix} 8-\lambda & -2 \\ -2 & 5-\lambda \end{vmatrix} = 0 \\ \Rightarrow \quad & (8-\lambda)(5-\lambda) - 4 = 0 \\ \Rightarrow \quad & \lambda^2 - 13\lambda + 36 = 0 \end{aligned}$$

Solution. We can rewrite the given equation in quadratic form as follows:

$$Q(x) = x^T Ax, \text{ where } A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}.$$

Then

$$\begin{aligned} & \begin{vmatrix} 8-\lambda & -2 \\ -2 & 5-\lambda \end{vmatrix} = 0 \\ \Rightarrow \quad & (8-\lambda)(5-\lambda) - 4 = 0 \\ \Rightarrow \quad & \lambda^2 - 13\lambda + 36 = 0 \\ \Rightarrow \quad & (\lambda - 9)(\lambda - 4) = 0. \end{aligned}$$

Solution. We can rewrite the given equation in quadratic form as follows:

$$Q(x) = x^T Ax, \text{ where } A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}.$$

Then

$$\begin{aligned} & \begin{vmatrix} 8-\lambda & -2 \\ -2 & 5-\lambda \end{vmatrix} = 0 \\ \Rightarrow \quad & (8-\lambda)(5-\lambda) - 4 = 0 \\ \Rightarrow \quad & \lambda^2 - 13\lambda + 36 = 0 \\ \Rightarrow \quad & (\lambda - 9)(\lambda - 4) = 0. \end{aligned}$$

Clearly, the eigenvalues of A are $\lambda_1 = 9, \lambda_2 = 4$.

Solution... $\underline{\underline{\lambda_1 = 9}}$

Solution... $\underline{\underline{\lambda_1 = 9}}$

We have

$$\begin{pmatrix} -1 & -2 \\ -2 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.$$

Solution... $\underline{\underline{\lambda_1 = 9}}$

We have
$$\begin{pmatrix} -1 & -2 \\ -2 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.$$

So, the eigenvector associated with λ_1 is
$$v_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

Solution... $\underline{\underline{\lambda_1 = 4}}$

We have
$$\begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.$$

Solution... $\underline{\underline{\lambda_1 = 4}}$

We have

$$\begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.$$

So, the eigenvector associated with λ_2 is

$$v_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Solution... $\underline{\underline{\lambda_1 = 4}}$

We have
$$\begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.$$

So, the eigenvector associated with λ_2 is
$$v_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Let $x = c_1 v_1 + c_2 v_2$, Now we have
$$Qx = x^T A x = x^T \lambda x = \lambda_1 c_1^2 + \lambda_2 c_2^2 = 9c_1^2 + 4c_2^2$$

Solution... $\underline{\underline{\lambda_1 = 4}}$

We have
$$\begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.$$

So, the eigenvector associated with λ_2 is
$$v_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Let $x = c_1 v_1 + c_2 v_2$, Now we have
$$Qx = x^T A x = x^T \lambda x = \lambda_1 c_1^2 + \lambda_2 c_2^2 = 9c_1^2 + 4c_2^2$$

Therefore, $Q(x) \geq 0$ and so $Q(0,0)$ is the global minimum.

Example 2. If $A = I_n$, then we get the following:

$$\begin{aligned} f(x) &= x_1^2 + x_2^2 + \ldots + x_n^2 \\ &= (x_1 \ \ldots \ x_n) \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{aligned}$$

Example 2. If $A = I_n$, then we get the following:

$$\begin{aligned} f(x) &= x_1^2 + x_2^2 + \ldots + x_n^2 \\ &= (x_1 \; \ldots \; x_n) \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= x^T I x = x^T x \end{aligned}$$

Example 2. If $A = I_n$, then we get the following:

$$\begin{aligned} f(x) &= x_1^2 + x_2^2 + ... + x_n^2 \\ &= \begin{pmatrix} x_1 & ... & x_n \end{pmatrix} \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= x^T I x = x^T x \\ &= \|x\|_2^2. \end{aligned}$$

Example 2. If $A = I_n$, then we get the following:

$$\begin{aligned} f(x) &= x_1^2 + x_2^2 + \dots + x_n^2 \\ &= \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= x^T I x = x^T x \\ &= \|x\|_2^2. \end{aligned}$$

This is the usual Euclidean norm of a vector $x \in \mathbb{R}^n$.

Example 3. If $A = D$, a diagonal matrix, e.g.,

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix},$$

then we get

$$f(x) = x_1^2 + 4x_2^2 + 9x_3^2$$

Example 3. If $A = D$, a diagonal matrix, e.g.,

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix},$$

then we get

$$\begin{aligned} f(x) &= x_1^2 + 4x_2^2 + 9x_3^2 \\ &= (x_1\ x_2\ x_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= x^T D x = (Dx^T)^{1/2} D^{1/2} x \end{aligned}$$

Example 3. If $A = D$, a diagonal matrix, e.g.,

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix},$$

then we get

$$\begin{aligned} f(x) &= x_1^2 + 4x_2^2 + 9x_3^2 \\ &= (x_1\ x_2\ x_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= x^T D x = (D x^T)^{1/2} D^{1/2} x \\ &= \|D^{1/2} x\|_2^2. \end{aligned}$$

Example 4. If A is a matrix that can be written as $A = B^T B$, i.e., it is a correlation/covariance matrix, then we get

$$\begin{aligned} f(x) &= x^T B^T B x \\ &= (Bx)^T Bx \\ &= \|Bx\|_2^2. \end{aligned}$$

Example 4. If A is a matrix that can be written as $A = B^T B$, i.e., it is a correlation/covariance matrix, then we get

$$\begin{aligned} f(x) &= x^T B^T B x \\ &= (Bx)^T Bx \\ &= \|Bx\|_2^2. \end{aligned}$$

Observe that a special case of this is the diagonal matrix with all positive entries that we saw in the previous example.

Problems. Consider the matrix

$$D = \begin{pmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3 \end{pmatrix}.$$

1. Show that if $\alpha_i > 0$, for all i, then the quadratic form function $f(x) = x^T Dx$ is a vector norm, in the sense that it satisfies the three conditions for a function to be a vector norm.

2. Show that if $\alpha_i \geq 0$, for alli, but $\alpha_i = 0$ for at least one i, then determine which of the conditions for the quadratic form function to be a norm are satisfied and which are violated.
3. Show that if $\alpha_i > 0$, for some i and $\alpha_i < 0$ for other i, then determine which of the conditions for the quadratic form function to be a norm are satisfied and which are violated.