Quadratic Forms

Definition 1: *Quadratic Form*

A **quadratic form** on $\mathbb{R}^N$ is a *polynomial* in the variables $x_1, x_2, \ldots, x_N$, so that all terms have *degree 2*. In other words, it is a function

$$q : \mathbb{R}^N \longrightarrow \mathbb{R}$$

of the form

$$\begin{array}{rcccccccccccccc} q(\mathbf{x}) & = & a_{11}x_1^2 & + & a_{12}x_1x_2 & + & a_{13}x_1x_3 & + & a_{14}x_1x_4 & + & \ldots & + & a_{1N}x_1x_N \\ & & & + & a_{22}x_2^2 & + & a_{23}x_2x_3 & + & a_{24}x_2x_4 & + & \ldots & + & a_{2N}x_2x_N \\ & & & & & + & a_{33}x_3^2 & + & a_{34}x_3x_4 & + & \ldots & + & a_{3N}x_3x_N \\ & & & & & & & \ddots & & \ddots & & \ddots & \ddots & \vdots & \vdots \\ & & & & & & & & & & & & + & a_{NN}x_N^2 \end{array}$$

where the $a_{i,j} \in \mathbb{R}$ are arbitrary constants.

Example 2:

- $f(x_1, x_2) = x_1^2 - 3x_1x_2 + 5x_2^2$ is a quadratic form on $\mathbb{R}^2$.
- $g(x_1, x_2, x_3) = x_1^2 - 7x_1x_2 + \frac{1}{3}x_1x_3 - \sqrt{2}x_2x_3 - 2x_2^2 + 6x_3^2$ is a quadratic form on $\mathbb{R}^3$.
- For any N, $|\mathbf{x}|^2 \;=\; x_1^2 + x_2^2 + \ldots + x_N^2$ is a quadratic form on $\mathbb{R}^N$.
- For any N, and any $a_1, a_2, \ldots, a_N \in \mathbb{R}$, $h(\mathbf{x}) \;=\; a_1x_1^2 + a_2x_2^2 + \ldots + a_Nx_N^2$ is a quadratic form on $\mathbb{R}^N$.

There is a close relationship between quadratic forms and $N \times N$ matrices as described by the following theorem.

Proposition 3: *Quadratic Forms and Matrices*

1. If $\boxed{\mathsf{A}}$ is any $N \times N$ matrix, then the function $p : \mathbb{R}^N \longrightarrow \mathbb{R}$ defined:

$$p(\mathbf{x}) \;=\; \mathbf{x}^t \cdot \boxed{\mathsf{A}} \cdot \mathbf{x}$$

is a quadratic form (where we regard $\mathbf{x} = \begin{bmatrix} \uparrow \\ \mathbf{x} \\ \downarrow \end{bmatrix}$ as a column-vector).

2. If $f : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ is a linear transformation, then the function $p : \mathbb{R}^N \longrightarrow \mathbb{R}$ defined:

$$p(\mathbf{x}) \;=\; \mathbf{x} \bullet f(\mathbf{x})$$

is a quadratic form.

3. If $q : \mathbb{R}^N \longrightarrow \mathbb{R}$ is *any* quadratic form, then there is a $N \times N$ matrix $\boxed{\mathsf{A}}$ so that $q(\mathbf{x}) \;=\; \mathbf{x}^t \cdot \boxed{\mathsf{A}} \cdot \mathbf{x}$, and a linear transformation f so that $q(\mathbf{x}) = \mathbf{x} \bullet f(\mathbf{x})$.

4. A matrix and its transpose determine the *same* quadratic form. In other words, if $\boxed{\mathsf{A}}$ is an $N \times N$ matrix, and we define quadratic forms $p(\mathbf{x}) = \mathbf{x}^t \cdot \boxed{\mathsf{A}} \cdot \mathbf{x}$ and $q(\mathbf{x}) = \mathbf{x}^t \cdot \boxed{\mathsf{A}}^t \mathbf{x}$, then $p(\mathbf{x}) = q(\mathbf{x})$.

5. Suppose $\boxed{\mathsf{A}_1}$ and $\boxed{\mathsf{A}_2}$ are $N \times N$ matrices, with $\boxed{\mathsf{A}} \;=\; \boxed{\mathsf{A}_1} + \boxed{\mathsf{A}_2}$. Suppose we define quadratic forms:

$$\begin{aligned} p_1(\mathbf{x}) &= \mathbf{x}^t \cdot \boxed{\mathsf{A}_1}\,\mathbf{x} \\ p_2(\mathbf{x}) &= \mathbf{x}^t \cdot \boxed{\mathsf{A}_2}\,\mathbf{x} \\ p(\mathbf{x}) &= \mathbf{x}^t \cdot \boxed{\mathsf{A}}\,\mathbf{x} \end{aligned}$$

Then $p(\mathbf{x}) = p_1(\mathbf{x}) + p_2(\mathbf{x})$.

6. There is a *unique* **symmetric** $N \times N$ matrix $\boxed{\mathsf{S}}$, so that $q(\mathbf{x}) \;=\; \mathbf{x}^t \cdot \boxed{\mathsf{S}} \cdot \mathbf{x}$. This is called the **matrix** of the form q.

Proof:

Proof of Part 1: Let $\boxed{\mathrm{A}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$.

Then

$$\begin{aligned} \mathbf{x}^t \cdot \boxed{\mathrm{A}} \cdot \mathbf{x} &= \begin{bmatrix} x_1 & \ldots & x_N \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \\ &= \begin{bmatrix} x_1 & \ldots & x_N \end{bmatrix} \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1N}x_N \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2N}x_N \\ \vdots \\ a_{N1}x_1 + a_{N,2}x_2 + \ldots + a_{N,N}x_N \end{bmatrix} \end{aligned}$$

$$= \sum_{n=1}^{N}\sum_{m=1}^{N} a_{nm}x_n x_m$$

Proof of Part 2: Let $\boxed{\mathrm{F}}$ be the standard matrix representation of f. Then $\mathbf{x} \bullet f(\mathbf{x}) = \mathbf{x}^t \cdot \boxed{\mathrm{A}} \cdot \mathbf{x}$.

Proof of Part 3: Suppose that

$$\begin{array}{rcccccccccccc} q(\mathbf{x}) & = & a_{11}x_1^2 & + & a_{12}x_1x_2 & + & a_{13}x_1x_3 & + & \dots & + & a_{1N}x_1x_N \\ & & & + & a_{22}x_2^2 & + & a_{23}x_2x_3 & + & \dots & + & a_{2N}x_2x_N \\ & & & & & + & a_{33}x_3^2 & + & \dots & + & a_{3N}x_3x_N \\ & & & & & & & \ddots & \ddots & \vdots & \vdots \\ & & & & & & & & & + & a_{NN}x_N^2 \end{array}$$

Then let $\boxed{\mathrm{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1,N} \\ 0 & a_{22} & a_{23} & \dots & a_{1,N} \\ 0 & 0 & a_{33} & \dots & a_{3,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{N,N} \end{bmatrix}$, and let $f : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be the corresponding linear transformation.

Proof of Part 4: **Exercise**. Hint: for any i and j, note that $x_i \cdot x_j = x_j \cdot x_i$.

Proof of Part 5: **Exercise**.

Proof of Part 6: **Existence:** Let $q : \mathbb{R}^N \longrightarrow \mathbb{R}$ be a quadratic form, and, by **Part 3**, find some $N \times N$ matrix $\boxed{\mathrm{A}}$ so that $q(\mathbf{x}) = \mathbf{x}^t \cdot \boxed{\mathrm{A}} \cdot \mathbf{x}$. Then by **Part 4**, we also have $q(\mathbf{x}) = \mathbf{x}^t \cdot \boxed{\mathrm{A}}^t \mathbf{x}$. Now define

$$\begin{array}{rcl} \boxed{\mathrm{S}} & = & \frac{1}{2}\boxed{\mathrm{A}} + \frac{1}{2}\boxed{\mathrm{A}}^t \\ \text{then } \mathbf{x}^t \cdot \boxed{\mathrm{S}}\mathbf{x} & = & \mathbf{x}^t \cdot \frac{1}{2}\boxed{\mathrm{A}} \cdot \mathbf{x} \;+\; \mathbf{x}^t \cdot \frac{1}{2}\boxed{\mathrm{A}}^t \cdot \mathbf{x} \\ & = & \frac{1}{2}\mathbf{x}^t \cdot \boxed{\mathrm{A}} \cdot \mathbf{x} \;+\; \frac{1}{2}\mathbf{x}^t \cdot \boxed{\mathrm{A}}^t \cdot \mathbf{x} \\ & = & \frac{1}{2}q(\mathbf{x}) \;+\; \frac{1}{2}q(\mathbf{x}) \\ & = & q(\mathbf{x}) \end{array}$$

Uniqueness: Suppose $\boxed{\mathsf{S}_1}$ and $\boxed{\mathsf{S}_2}$ are two symmetric, $N \times N$ matrices so that

$$\mathbf{x}^t \cdot \boxed{\mathsf{S}_1} \cdot \mathbf{x} \;=\; q(\mathbf{x}) \;=\; \mathbf{x}^t \cdot \boxed{\mathsf{S}_2} \cdot \mathbf{x}.$$

We want to show that $\boxed{\mathsf{S}_1} = \boxed{\mathsf{S}_2}$. Thus, if $\boxed{\mathsf{Z}} = \boxed{\mathsf{S}_1} - \boxed{\mathsf{S}_2}$, we want to show $\boxed{\mathsf{Z}} = 0$. But for any $\mathbf{x} \in \mathbb{R}^N$,

$$
\begin{aligned}
\mathbf{x}^t \cdot \boxed{\mathsf{Z}}\mathbf{x} &= \mathbf{x}^t \cdot \left(\boxed{\mathsf{S}_1} - \boxed{\mathsf{S}_2}\right) \cdot \mathbf{x} \\
&= \mathbf{x}^t \cdot \boxed{\mathsf{S}_1} \cdot \mathbf{x} \quad - \quad \mathbf{x}^t \cdot \boxed{\mathsf{S}_2} \cdot \mathbf{x} \\
&= q(\mathbf{x}) - q(\mathbf{x}) \\
&= 0 \qquad\qquad (A)
\end{aligned}
$$

In particular, for any i, j, let

$$
\mathbf{e}_i = \left(\overbrace{\underbrace{0, \ldots, 0}_{(i-1)}, 1, 0, \ldots, 0}^{N} \right) \qquad \text{and} \quad \mathbf{e}_j = \left(\overbrace{\underbrace{0, \ldots, 0}_{(j-1)}, 1, 0, \ldots, 0}^{N} \right)
$$

be standard basis vectors, and let $\mathbf{x} = \mathbf{e}_i + \mathbf{e}_j$. Then

$$
\begin{aligned}
0 &=_{(1)} (\mathbf{e}_i + \mathbf{e}_j)^t \cdot \boxed{\mathsf{Z}} \cdot (\mathbf{e}_i + \mathbf{e}_j) \\
&= \mathbf{e}_i^t \cdot \boxed{\mathsf{Z}} \cdot \mathbf{e}_i \; + \; \mathbf{e}_i^t \cdot \boxed{\mathsf{Z}} \cdot \mathbf{e}_j \; + \; \mathbf{e}_j^t \cdot \boxed{\mathsf{Z}} \cdot \mathbf{e}_i \; + \; \mathbf{e}_j^t \cdot \boxed{\mathsf{Z}} \cdot \mathbf{e}_j \\
&=_{(2)} \mathbf{e}_i^t \cdot \boxed{\mathsf{Z}} \cdot \mathbf{e}_j \; + \; \mathbf{e}_j^t \cdot \boxed{\mathsf{Z}} \cdot \mathbf{e}_i \\
&=_{(3)} \mathbf{e}_i^t \cdot \boxed{\mathsf{Z}} \cdot \mathbf{e}_j \; + \; \mathbf{e}_j^t \cdot \boxed{\mathsf{Z}}^t \cdot \mathbf{e}_i \\
&= \mathbf{e}_i^t \cdot \boxed{\mathsf{Z}} \cdot \mathbf{e}_j \; + \; \left(\mathbf{e}_i^t \cdot \boxed{\mathsf{Z}} \cdot \mathbf{e}_j\right)^t \\
&=_{(4)} \mathbf{e}_i^t \cdot \boxed{\mathsf{Z}} \cdot \mathbf{e}_j \; + \; \mathbf{e}_i^t \cdot \boxed{\mathsf{Z}} \cdot \mathbf{e}_j \\
&= 2 \cdot \mathbf{e}_i^t \cdot \boxed{\mathsf{Z}} \cdot \mathbf{e}_j \\
&=_{(5)} 2 \cdot z_{ij}
\end{aligned}
$$

(1) and (2) follow from Equation (A).

(3) Because $\boxed{\mathsf{Z}}$ is symmetric

(4) A scalar is a 1×1 matrix, so it is automatically equal to its transpose.

(5) Where z_{ij} is the (i, j)th element of $\boxed{\mathsf{Z}}$.

Hence, we conclude: for all i, j, $z_{ij} = 0$. In other words, $\boxed{\mathsf{Z}} = 0$; in other words $\boxed{\mathsf{S}_1} = \boxed{\mathsf{S}_2}$.

□ [Proposition 3]

Example 4:

- If $f(x_1, x_2) = x_1^2 - 3x_1x_2 + 5x_2^2$, then

$$\begin{aligned} f(x_1,x_2) &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \cdot \begin{bmatrix} 1 & -3 \\ 0 & 5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \\ \text{But also, } f(x_1,x_2) &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ -3 & 5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \\ \text{and, } f(x_1,x_2) &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \cdot \begin{bmatrix} 1 & -3/2 \\ -3/2 & 5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \end{aligned}$$

  This last matrix is the unique *symmetric* matrix for f.

- If $g(x_1, x_2, x_3) = x_1^2 - 7x_1x_2 + \frac{1}{3}x_1x_3 - \sqrt{2}x_2x_3 - 2x_2^2 + 6x_3^2$, then

$$g(x_1,x_2,x_3) = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \cdot \begin{bmatrix} 1 & -\frac{7}{2} & \frac{1}{6} \\ -\frac{7}{2} & -2 & -\frac{\sqrt{2}}{2} \\ \frac{1}{6} & -\frac{\sqrt{2}}{2} & 6 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

- If $|\mathbf{x}|^2 = x_1^2 + x_2^2 + \ldots + x_N^2$, then

$$|\mathbf{x}|^2 = \begin{bmatrix} x_1 & \ldots & x_N \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{bmatrix}$$

- If $h(\mathbf{x}) = a_1x_1^2 + a_2x_2^2 + \ldots + a_Nx_N^2$, then

$$h(\mathbf{x}) = \begin{bmatrix} x_1 & \ldots & x_N \end{bmatrix} \cdot \begin{bmatrix} a_1 & 0 & 0 & \ldots & 0 \\ 0 & a_2 & 0 & \ldots & 0 \\ 0 & 0 & a_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & a_N \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{bmatrix}.$$

Theorem 5: *Principal Axis Theorem for Quadratic Forms*
Let $q : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be a quadratic form. Then

1. If $\boxed{\mathsf{Q}}$ is a symmetric matrix so that $q(\mathbf{x}) = \mathbf{x}^t \cdot \boxed{\mathsf{Q}} \cdot \mathbf{x}$, then there is an **orthogonal matrix** $\boxed{\mathsf{F}}$ which diagonalizes $\boxed{\mathsf{Q}}$ (ie. $\boxed{\mathsf{D}} = \boxed{\mathsf{F}} \cdot \boxed{\mathsf{Q}} \cdot \boxed{\mathsf{F}}^t$ is diagonal), and

$$q(\mathbf{x}) = \left(\boxed{\mathsf{F}} \cdot \mathbf{x}\right)^t \cdot \boxed{\mathsf{D}} \cdot \left(\boxed{\mathsf{F}} \cdot \mathbf{x}\right) \qquad (A)$$

2. There is an orthogonal transformation $f : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ and a **diagonal** matrix $\boxed{\mathsf{D}}$ so that, if we define the new quadratic form

$$p(\mathbf{x}) = \mathbf{x}^t \cdot \boxed{\mathsf{D}} \cdot \mathbf{x}$$

then $q(\mathbf{x}) = p\left(f(\mathbf{x})\right)$ $\qquad (B)$.

3. There is an orthonormal basis $\widetilde{\mathcal{B}} = \{\mathbf{b}_1, \ldots, \mathbf{b}_N\}$ and real numbers $d_1, \ldots, d_N \in \mathbb{R}$ so that, for any $\mathbf{x} \in \mathbb{R}^D$, if $\mathbf{x}$ has coordinate N-tuple $(\widetilde{x}_1, \ldots, \widetilde{x}_N)$ relative to $\widetilde{\mathcal{B}}$, then

$$q(\mathbf{x}) = d_1\widetilde{x}_1^2 + d_2\widetilde{x}_2^2 + \ldots d_N\widetilde{x}_N^2 \qquad (C)$$

Proof:

Proof of Part 1: By the Spectral Theorem for Symmetric Matrices, we can find an orthogonal matrix $\boxed{\mathrm{F}}$ diagonalizing $\boxed{\mathrm{Q}}$. If $\boxed{\mathrm{D}} = \boxed{\mathrm{F}} \cdot \boxed{\mathrm{Q}} \cdot \boxed{\mathrm{F}}^t$, then $\boxed{\mathrm{Q}} = \boxed{\mathrm{F}}^t \cdot \boxed{\mathrm{D}} \cdot \boxed{\mathrm{F}}$, and therefor,

$$\begin{aligned} q(\mathbf{x}) &= \mathbf{x}^t \cdot \boxed{\mathrm{Q}} \cdot \mathbf{x} \\ &= \mathbf{x}^t \cdot \boxed{\mathrm{F}}^t \cdot \boxed{\mathrm{D}} \cdot \boxed{\mathrm{F}} \cdot \mathbf{x} \\ &= \left(\boxed{\mathrm{F}} \cdot \mathbf{x}\right)^t \cdot \boxed{\mathrm{D}} \cdot \left(\boxed{\mathrm{F}} \cdot \mathbf{x}\right) \end{aligned}$$

Proof of Part 2: Let $f : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be the orthogonal transformation determined by the matrix $\boxed{\mathrm{F}}$ from **Part 1**. Then equation (B) is just a restatement of equation (A).

Proof of Part 3: Let $\widetilde{\mathcal{B}}$ be the orthonormal basis for $\mathbb{R}^N$ such that $\boxed{\mathrm{F}}$ is the change-of-basis matrix from the standard basis into $\widetilde{\mathcal{B}}$. Let $\boxed{\mathrm{D}} = \begin{bmatrix} d_1 & 0 & 0 & \ldots & 0 \\ 0 & d_2 & 0 & \ldots & 0 \\ 0 & 0 & d_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & d_N \end{bmatrix}$. Then equation (C) is just a restatement of equation (A).

$\square$ [Theorem 5]

Definition 6: *Principle Axes*

Let $q : \mathbb{R}^D \longrightarrow \mathbb{R}$ be a quadratic form, and let $\widetilde{\mathcal{B}}$ be the orthonormal basis described in **Part 3** of the previous theorem. The elements of $\widetilde{\mathcal{B}}$ are called the **principal axes** of q.

Remark 7: If $\boxed{\mathrm{Q}}$ is the unique symmetric matrix so that $q(\mathbf{x}) = \mathbf{x}^t \cdot \boxed{\mathrm{Q}} \cdot \mathbf{x}$, then the principal axes of q constitute an *orthonormal basis* for $\mathbb{R}^D$ consisting of *eigenvectors* of $\boxed{\mathrm{Q}}$.

Proof: Exercise

$\square$ [Remark 7]

Example 8: *(plagiarised from Nicholson)*

$$\begin{aligned}
\text{Let } q(\mathbf{x}) &= 3\left(x_1^2 + x_2^2 + x_3^2 + x_4^2\right) + 2\left(x_1 x_2 + x_3 x_4\right) \\
&\qquad + 10\left(x_1 x_4 + x_2 x_3 - x_1 x_3 - x_2 x_4\right) \\
&= \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix} \begin{bmatrix} 3 & 1 & -5 & 5 \\ 1 & 3 & 5 & -5 \\ -5 & 5 & 3 & 1 \\ 5 & -5 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \\
&= \mathbf{x}^t \cdot \boxed{\mathrm{A}} \cdot \mathbf{x}
\end{aligned}$$

$\boxed{\mathrm{A}}$ has eigenvectors

$$\mathbf{b}_1 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{b}_2 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{b}_3 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \text{and } \mathbf{b}_4 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}.$$

with corresponding eigenvalues

$$\lambda_1 = 12, \qquad \lambda_2 = -8, \qquad \lambda_3 = \lambda_4 = 4.$$

$$\text{So, if we define } \boxed{\mathrm{F}} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix},$$

$$\begin{aligned}
\text{then } \boxed{F}^t \cdot \boxed{A} \cdot \boxed{F} &= \begin{bmatrix} 12 &&& \\ & -8 && \\ && 4 & \\ &&& 4 \end{bmatrix}, \\
\text{and if we define } \mathbf{y} &= \boxed{F} \cdot \mathbf{x} \\
\text{then } \quad q(\mathbf{x}) &= p(\mathbf{y}) \\
\text{where } p(\mathbf{y}) &= 12y_1^2 - 8y_2^2 + 4(y_3^2 + y_4^2).
\end{aligned}$$

Definition 9: *Index, Rank*

Let $q : \mathbb{R}^D \longrightarrow \mathbb{R}$ be a quadratic form and let $\boxed{Q}$ be the symmetric matrix so that $q(\mathbf{x}) = \mathbf{x}^t \cdot \boxed{Q} \cdot \mathbf{x}$.

The **index** of $\boxed{Q}$ is the number of *positive eigenvalues* of $\boxed{Q}$. The **index** of q is the index of $\boxed{Q}$.

The **rank** of q is the rank of the matrix $\boxed{Q}$.

Remark 10: Suppose that $\boxed{B}$ is the change-of-basis matrix from the principal axes basis for q to the standard basis $\mathcal{E}$. Thus, if $\boxed{D} = \boxed{B}^t \cdot \boxed{A} \cdot \boxed{B}$, then $\boxed{D}$ will be a diagonal matrix, as described in Theorem 5 on page 5. By suitable "reordering" of the principal axes, we can ensure that $\boxed{D}$ is of the form:

$$\boxed{D} = \left[\begin{array}{c|c|c}
\begin{matrix} +\lambda_1 &&& \\ & +\lambda_2 && \\ && \ddots & \\ &&& +\lambda_P \end{matrix} & & \\ \hline
& \begin{matrix} -\kappa_1 &&& \\ & -\kappa_2 && \\ && \ddots & \\ &&& -\kappa_N \end{matrix} & \\ \hline
& & \underbrace{\begin{matrix} 0 &&& \\ & 0 && \\ && \ddots & \\ &&& 0 \end{matrix}}_{Z}
\end{array}\right]$$

where

- All blank entries indicate **zeros**.
- $+\lambda_1, \ldots, +\lambda_P$ are **positive**.
- $-\kappa_1, \ldots, -\kappa_N$ are **negative**.
- P is the **index** of q,
- $P+N$ is the **rank** of q,
- $P+N+Z = D$.

Remark 11: *"Complete" Diagonalization*
As in the previous remark, let $\widetilde{\mathcal{B}} = \{\mathbf{b}_1, \ldots, \mathbf{b}_D\}$ be the orthonormal basis of principal axes for the quadratic form q. Now, define a *new* basis $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_D\}$ as follows:

$$\begin{array}{rclcrcl}
\mathbf{c}_1 & = & \sqrt{\lambda_1} \cdot \mathbf{b}_1, & \ldots & \mathbf{c}_P & = & \sqrt{\lambda_P} \cdot \mathbf{b}_P; \\
\mathbf{c}_{P+1} & = & \sqrt{\kappa_1} \cdot \mathbf{b}_{P+1}, & \ldots & \mathbf{c}_{P+N} & = & \sqrt{\kappa_N} \cdot \mathbf{b}_{P+N}; \\
\mathbf{c}_{P+N+1} & = & \mathbf{b}_{P+N+1}, & \ldots & \mathbf{c}_D & = & \mathbf{b}_D;
\end{array}$$

Again, let $\boxed{\mathrm{C}}$ be the change-of-basis matrix from the principal axes basis for q to the standard basis $\mathcal{E}$. Thus, if $\boxed{\mathrm{D}} = \boxed{\mathrm{C}}^t \cdot \boxed{\mathrm{A}} \cdot \boxed{\mathrm{C}}$, then $\boxed{\mathrm{D}}$ will be a diagonal matrix of the form

$$\boxed{\mathrm{D}} = \left[\begin{array}{cccc|cccc|cccc}
+1 & & & & & & & & & & & \\
 & +1 & & & & & & & & & & \\
 & & \ddots & & & & & & & & & \\
 & & & +1 & & & & & & & & \\
\hline
 & & & & -1 & & & & & & & \\
 & & & & & -1 & & & & & & \\
 & & & & & & \ddots & & & & & \\
 & & & & & & & -1 & & & & \\
\hline
 & & & & & & & & 0 & & & \\
 & & & & & & & & & 0 & & \\
 & & & & & & & & & & \ddots & \\
 & & & & & & & & \underbrace{ }_{} & & \underbrace{}_{Z} & 0
\end{array}\right]$$

This procedure is called **completely diagonalizing** the quadratic form q.

Definition 12: *Signature*

The triple (P, N, Z) is called the **signature** of the quadratic form.

Definition 13: *Positive Definite*

Let $q : \mathbb{R}^D \longrightarrow \mathbb{R}$ be a quadratic form. q is called **positive definite** if, for *every* $\mathbf{x} \in \mathbb{R}^D$, with $\mathbf{x} \neq 0$, $\quad q(\mathbf{x}) > 0$.

Example 14: Suppose $q(\mathbf{x}) = \mathbf{x}^t \cdot \boxed{\mathrm{D}} \cdot \mathbf{x}$, where $\boxed{\mathrm{D}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_D \end{bmatrix}$.

$$\text{Then} \quad \left(q \text{ is } \textbf{positive definite} \right) \iff \left(\lambda_1, \lambda_2, \dots, \lambda_D > 0 \right).$$

This example illustrates a general principle....

Proposition 15: Let $q : \mathbb{R}^D \longrightarrow \mathbb{R}$ be a quadratic form. The following are equivalent:

- q is **positive definite**.
- The **index** of q is D.
- The **signature** of q is $(D, 0, 0)$.

Proof: Exercise.

□ [Proposition 15]

Application: Conics and Quadrics (optional)[1]

Definition 16: *Conic*

Let $q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a quadratic form, let $\mathbf{v} \in \mathbb{R}^2$ be some vector, and $r \in \mathbb{R}$ some scalar. The set

$$\mathbb{S}_r = \left\{ \mathbf{x} \in \mathbb{R}^2 \; ; \; q(\mathbf{x}) + \mathbf{x} \bullet \mathbf{v} = r \right\}$$

is called an **conic**. If the term $\mathbf{v}$ is zero, the conic is called **homogeneous**.

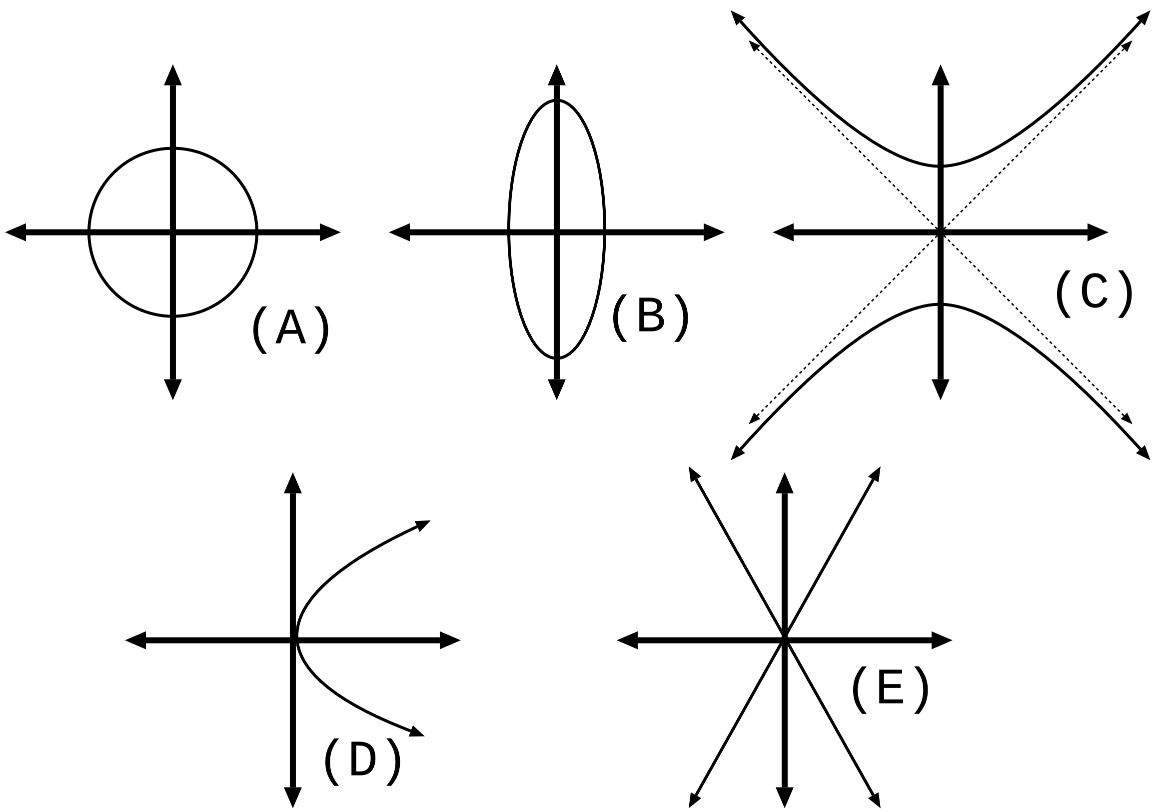

Figure 1: The five classes of conics

Remark 17: There are five basic classes of conics....

- **Circles**, generated by equations of the form $x_1^2 + x_2^2 = R$, where $R > 0$. (**Part A** of Figure 1 on the page before)
- **Ellipses**, such as those generated by equations of the form $a_1 x_1^2 + a_2 x_2^2 = R$, for some $a_1, a_2 > 0$. where $R > 0$. (Thus, really, a **circle** is just a special kind of ellipse, where $a_1 = a_2 = 1$. (**Part B** of Figure 1 on the preceding page)
- **Hyperbolas**, such as those generated by equations of the form $a_1 x_1^2 - a_2 x_2^2 = R$, for some $a_1, a_2 > 0$, and $R \in \mathbb{R}$. (**Part C** of Figure 1 on the page before)
- **Parabolas**, such as those generated by equations of the form $a_1 x_1^2 + b_2 x_2 = R$, for nonzero $a_1, b_2 \in \mathbb{R}$ and arbitrary $R \in \mathbb{R}$. (**Part D** of Figure 1 on the preceding page)
- Anything else is called a **Degenerate Conic**, such as those generated by equations of the form $x_1^2 - x_2^2 = 0$ (ie. $R = 0$.) (**Part E** of Figure 1 on the page before)

The previous examples are just the special cases when the "principal axes" of the conic coincide with the standard axes of $\mathbb{R}^2$. The actual definitions are as follows:

Definition 18: *Ellipse*

Let $q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a *positive-definite* quadratic form. For any $r > 0$, the set of the form

$$\mathbb{E} \;=\; \left\{\mathbf{x} \in \mathbb{R}^2 \;;\; q(\mathbf{x}) = r\right\}$$

is called an **ellipse** (centered at the origin).

Definition 19: *Hyperbola*

Let $q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a quadratic form with *index* 1 and *rank* 2. For any $r \in \mathbb{R}$, the set of the form

$$\mathbb{H} \;=\; \left\{\mathbf{x} \in \mathbb{R}^2 \;;\; q(\mathbf{x}) = r\right\}$$

is called a **hyperbola** (centered at the origin).

[1] For simplicity, in this section we consider only conics centered at the origin.

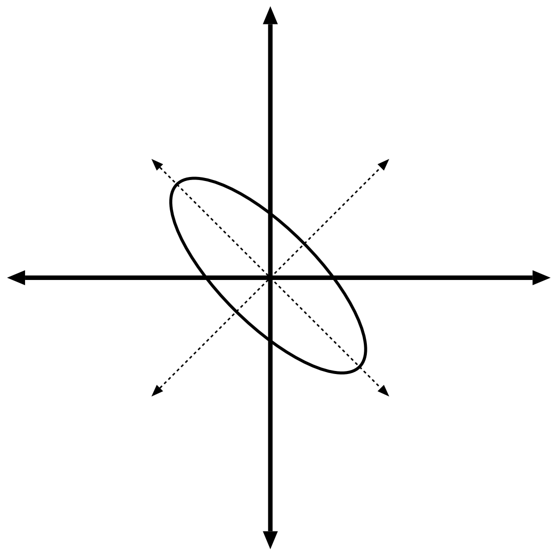

Figure 2: The ellipse $3x_1^2 - x_1x_2 + 3x_2^2 = 2$

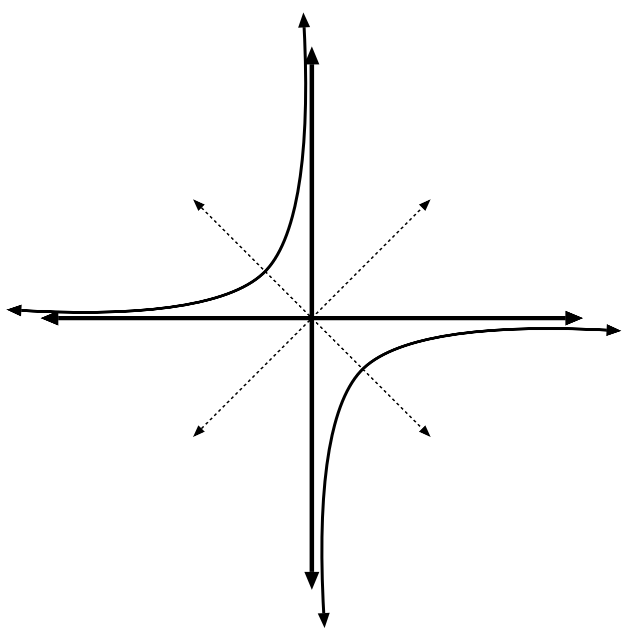

Figure 3: The Hyperbola $x_1x_2 = 1$

Definition 20: *Parabola*

Let $q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a quadratic form with *rank* 1. For any $\mathbf{v} \in \mathbb{R}^2$, the set of the form

$$\mathbb{P} \;=\; \left\{\mathbf{x} \in \mathbb{R}^2 \;;\; q(\mathbf{x}) = \mathbf{x} \bullet \mathbf{v}\right\}$$

is called a **parabola** (centered at the origin).

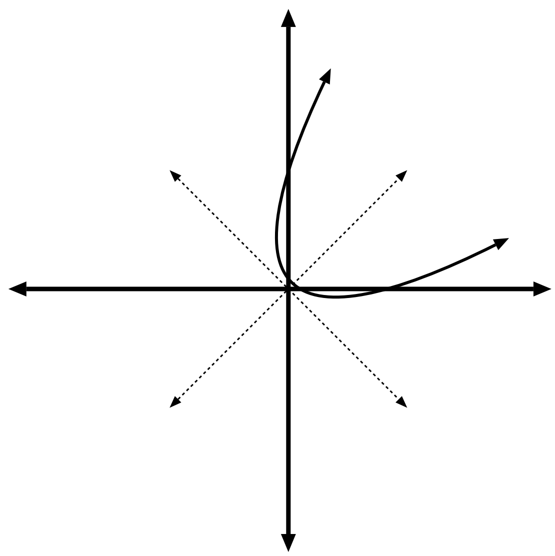

Figure 4: The parabola $x_1^2 - 2x_1x_2 + x_2^2 \;=\; x_1 + x_2$

Definition 21: *Degenerate Conic*

Let $q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a quadratic form with *rank 0*. For any $r \in \mathbb{R}$, the set of the form

$$\mathbb{D} \;=\; \left\{\mathbf{x} \in \mathbb{R}^2 \;;\; q(\mathbf{x}) = r\right\}$$

is called a **degenerate conic**.

The previous examples suggest that an arbitrary conic can be obtained from one of the simple examples given in Remark 17 on page 12 through some rotation of the plane.

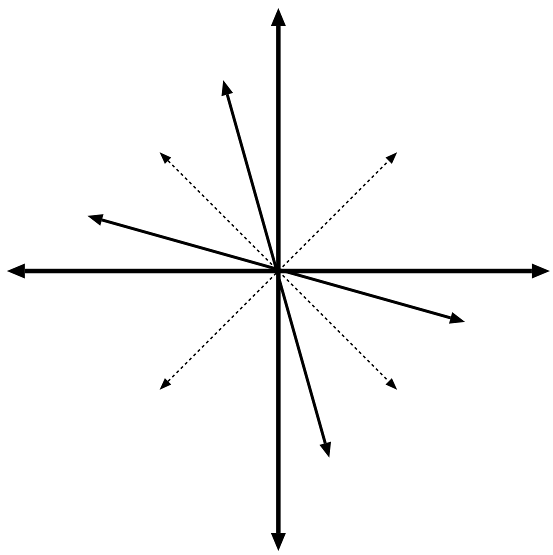

Figure 5: A degenerate conic.

Proposition 22: *Principle Axis Theorem For Homogeneous Conics* Let $q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be any quadratic form, let $R \in \mathbb{R}$, and let $\mathbb{S}_r = \{\mathbf{x} \in \mathbb{R}^2 \ ; \ q(\mathbf{x}) = R\}$ be a homogeneous conic. Then there is an orthogonal transformation $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ (ie. a rotation and/or reflection of the plane) so that $T(\mathbb{S}_r)$ has the form of one of the standard examples of Remark 17 on page 12.

Proof: Exercise. Use the **Principle Axis Theorem for Quadrics** described below (Theorem 27 on the following page).

□ [Proposition 22]

The generalizations of conics to higher dimensions are called **quadrics**....

Definition 23: *Quadric*

Let $q : \mathbb{R}^D \longrightarrow \mathbb{R}$ be a quadratic form, $\mathbf{v} \in \mathbb{R}^D$ a vector, and $r \in \mathbb{R}$ a scalar. The set

$$\mathbb{S}_r = \{\mathbf{x} \in \mathbb{R}^D \ ; \ q(\mathbf{x}) + \mathbf{v} \bullet \mathbf{x} = r\}$$

is called an ($(D-1)$-dimensional) **quadric**. If $\mathbf{v} = 0$, the quadric is called **homogeneous**.

Definition 24: *Ellipsoid*

Let $q : \mathbb{R}^D \longrightarrow \mathbb{R}$ be a quadratic form whose *index* and *rank* are both equal to D. For any $r > 0$, the set of the form

$$\mathbb{E} \ = \ \left\{\mathbf{x} \in \mathbb{R}^D \ ; \ q(\mathbf{x}) = r\right\}$$

is called a **ellipsoid**.

Definition 25: *Hyperboloid*

Let $q : \mathbb{R}^D \longrightarrow \mathbb{R}$ be a quadratic form of rank D, whose *index* is less than D and bigger than 1. For any $r \in \mathbb{R}$, the set of the form

$$\mathbb{H} \ = \ \left\{\mathbf{x} \in \mathbb{R}^D \ ; \ q(\mathbf{x}) = r\right\}$$

is called a **hyperboloid**

Definition 26: *Paraboloid*

Let $q : \mathbb{R}^D \longrightarrow \mathbb{R}$ be a quadratic form, and let $\mathbf{v} \in \mathbb{R}^D$ be nonzero. For any $r \in \mathbb{R}$, the set of the form

$$\mathbb{P} \ = \ \left\{\mathbf{x} \in \mathbb{R}^D \ ; \ q(\mathbf{x}) + \mathbf{v} \bullet \mathbf{x} \ = \ r\right\}$$

is called a **paraboloid**. If the index of q is D or 1, then $\mathbb{P}$ is called an **elliptical paraboloid**; if the index is less than D and greater than 1, then $\mathbb{P}$ then $\mathbb{P}$ is called an **hyperbolic paraboloid**.

Proposition 27: *Principle Axis Theorem For Homogeneous Quadrics*
Let $q : \mathbb{R}^D \longrightarrow \mathbb{R}$ be any quadratic form, let $R \in \mathbb{R}$, and let $\mathbb{S}_r \ = \ \left\{\mathbf{x} \in \mathbb{R}^D \ ; \ q(\mathbf{x}) = R\right\}$ be a homogeneous quadric. Then there is an orthogonal transformation $T : \mathbb{R}^D \longrightarrow \mathbb{R}^D$ so that $T(\mathbb{S}_r)$ is a "standard" quadric; in other words,

$$T(\mathbb{S}_r) \ = \ \left\{\mathbf{x} \in \mathbb{R}^D; \quad \begin{array}{ccccccccc} \lambda_1 x_1^2 & + & \lambda_2 x_2^2 & + & \ldots & + & \lambda_P x_P^2 & & \\ -\kappa_1 x_{P+1}^2 & - & \kappa_2 x_{P+2}^2 & - & \ldots & - & \kappa_P x_{P+N}^2 & = & 0 \end{array}\right\}$$

where

- (P, N, Z) is the **signature** of the quadratic form q.
- $+\lambda_1, \ldots, +\lambda_N > 0$ are the *positive* eigenvalues of the matrix $\boxed{\mathsf{Q}}$ associated to q.

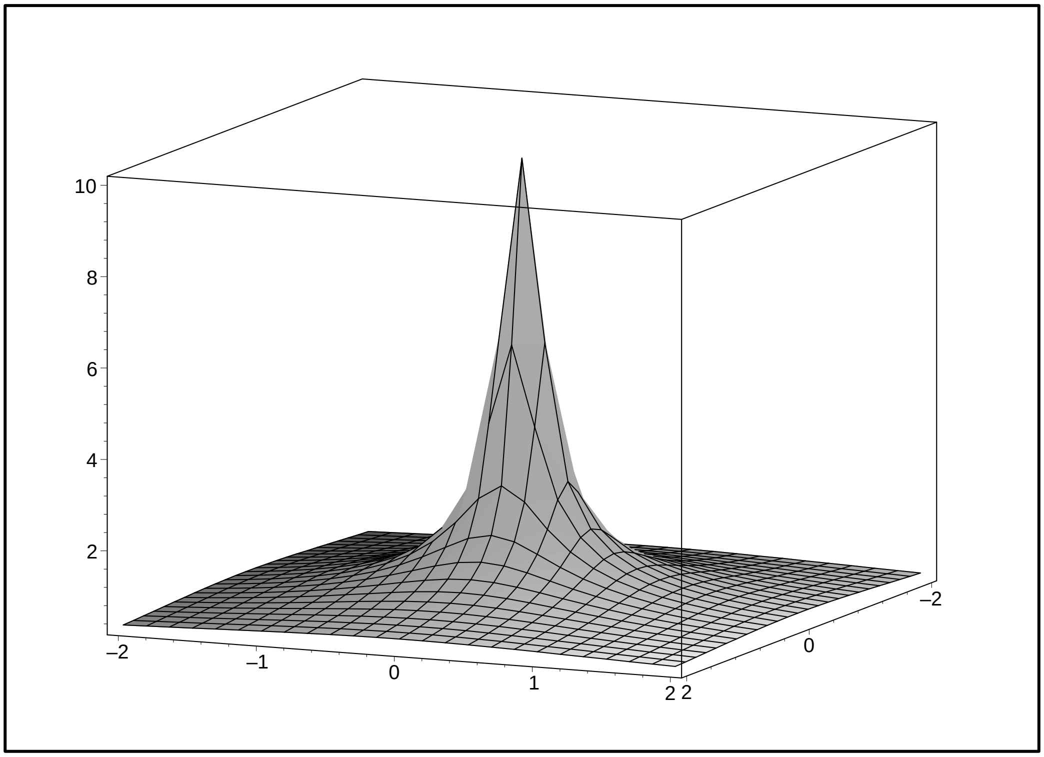

Figure 6: A hyperboloid in $\mathbb{R}^3$, with equation $x_1^2 + x_2^2 - x_3^2 = 1$

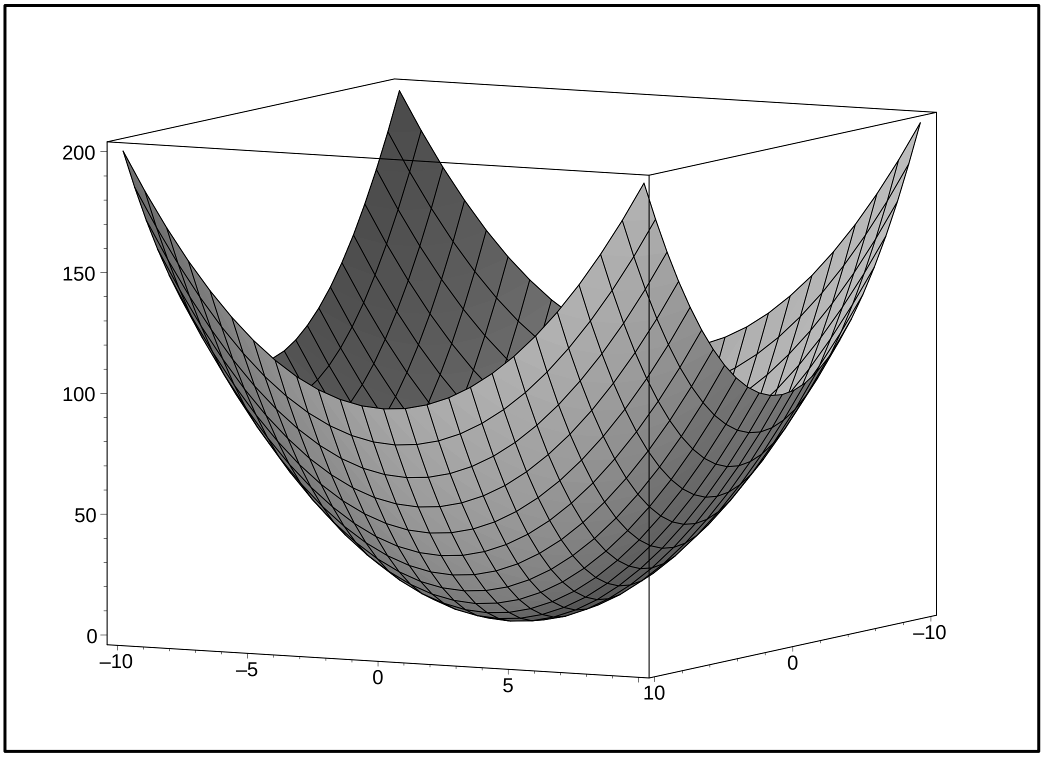

Figure 7: The elliptical paraboloid in $\mathbb{R}^3$, with equation $x_1^2 + x_2^2 = x_3$

- $-\kappa_1, \ldots, -\kappa_N < 0$ are the *negative* eigenvalues of $\boxed{\mathsf{Q}}$.

Proof: Exercise. Use the Principal Axis Theorem for Quadratic Forms (Theorem 5 on page 5).

□ [Proposition 27]

Remark 28: There is a corresponding version of this theorem for *nonhomogenous* quadrics, but more algebra is required to prove it.

Definition 29: *Principal Axes*

Let $\mathbb{S}_R \subset \mathbb{R}^D$ be a homogeneous quadric, induced by quadratic form $q : \mathbb{R}^D \longrightarrow \mathbb{R}$. Let $T : \mathbb{R}^D \longrightarrow \mathbb{R}^D$ be the orthogonal transformation described in the previous theorem. Let $\mathcal{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_D\}$ be the standard orthonormal basis of $\mathbb{R}^D$, and let $\mathbf{b}_d = T(\mathbf{e}_d)$ for all $d \in [1..D]$.

The vectors $\{\mathbf{b}_1, \ldots, \mathbf{b}_D\}$ are called the **principal axes** of $\mathbb{S}_R$.

Application: Normal Probability Distributions (optional)

Definition 30: *Univariate Normal Probability Distribution*

A (univariate) **normal** probability distribution is a function $\rho : \mathbb{R} \longrightarrow \mathbb{R}$ of the form:

$$\rho(x) \;=\; \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\left(\frac{x}{\sigma}\right)^2\right],$$

where $\sigma > 0$ is some constant.

This function represents a *probability density* on $\mathbb{R}$. The parameter $\sigma > 0$ is called the **standard deviation** of the distribution, and represents its "width"; the smaller σ is, the more "sharply peaked" the distribution becomes. The factor $\frac{1}{\sqrt{2\pi}\sigma}$ is present just to make sure that $\int_{\mathbb{R}} \rho \;=\; 1$.

Example 31: *Standard Normal Distribution*
For example, if $\sigma = 1$, we get the **standard** univariate normal distribution:

$$\rho(x) \;=\; \frac{1}{\sqrt{2\pi}} \exp\left[-x^2\right].$$

Definition 32: *Multivariate Normal Probability Distribution*

A **multivariate normal distribution** is a function $\rho : \mathbb{R}^N \longrightarrow \mathbb{R}$ of the form:

$$\rho(\mathbf{x}) \;=\; \sqrt{\frac{\det \boxed{\mathsf{A}}}{(2\pi)^N}} \exp\left[-\mathbf{x}^t \cdot \boxed{\mathsf{A}} \cdot \mathbf{x}\right],$$

where $\boxed{\mathsf{A}}$ is a symmetric, $N \times N$ *positive definite* matrix.

Now ρ represents a probability density on $\mathbb{R}^N$.

The condition that $\boxed{\mathsf{A}}$ be *symmetric* is not necessary for the definition to make sense. However, the expression $\mathbf{x}^t \cdot \boxed{\mathsf{A}} \cdot \mathbf{x}$ is a quadratic form, and thus, we can always assume that $\boxed{\mathsf{A}}$ is symmetric —indeed, for any multivariate normal distribution ρ, there is a *unique* symmetric matrix satisfying the formula given above.

The symmetric matrix $\boxed{\mathsf{C}} := \boxed{\mathsf{A}}^{-1}$ is called the **covariance matrix** of the distribution. Intuitively, the diagonal terms $c_{11}, c_{22}, \ldots, c_{NN}$ represent the "width" of the distribution along the dimensions $x_1, x_2, \ldots, x_N$, while cross-terms of the form $c_{i,j}$ represent the degree of "correlation" between the random variables x_i and x_j.

The condition that $\boxed{\mathsf{A}}$ be *positive definite* is analogous to the condition that the standard deviation σ be *positive* in the univariate case.

Example 33: *Standard Multivariate Normal Distribution*
For example, if $\boxed{\mathrm{A}} = \boxed{\mathbf{Id}}$, then

$$\mathbf{x}^t \cdot \boxed{\mathrm{A}} \cdot \mathbf{x} \;=\; \mathbf{x}^t \cdot \mathbf{x} \;=\; \mathbf{x} \bullet \mathbf{x} \;=\; x_1^2 + x_2^2 + \ldots + x_N^2,$$

and we get the **standard** multivariate normal distribution:

$$\begin{aligned}
\rho(\mathbf{x}) &= \frac{1}{(2\pi)^{N/2}} \exp\left[-\left(x_1^2 + x_2^2 + \ldots + x_N^2\right)\right] \\
&= \frac{1}{(2\pi)^{N/2}} \exp\left[-x_1^2 - x_2^2 - \ldots - x_N^2\right] \\
&= \frac{1}{(2\pi)^{N/2}} \cdot \exp\left[-x_1^2\right] \cdot \exp\left[-x_2^2\right] \cdot \ldots \cdot \exp\left[-x_N^2\right] \\
&= \frac{1}{\sqrt{2\pi}} \cdot \exp\left[-x_1^2\right] \cdot \frac{1}{\sqrt{2\pi}} \cdot \exp\left[-x_2^2\right] \cdot \ldots \cdot \frac{1}{\sqrt{2\pi}} \cdot \exp\left[-x_N^2\right] \\
&= \rho_1(x_1) \cdot \rho_2(x_2) \cdot \ldots \cdot \rho_N(x_N)
\end{aligned}$$

In other words, the standard normal distribution in $\mathbb{R}^N$ is a *product* of N independent *univariate* standard normal distributions in the variables $x_1, \ldots, x_N$. In the language of probability theory, we say that $x_1, \ldots, x_N$ are *independent* random variables, each having a standard normal distribution.

Example 34: *A Product of Univariate Distributions*

More generally, suppose if $\boxed{\mathrm{A}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ 0 & 0 & \lambda_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_N \end{bmatrix}$, with $\lambda_1, \lambda_2, \ldots, \lambda_N > 0$. Then:

$$\mathbf{x}^t \cdot \boxed{\mathrm{A}} \cdot \mathbf{x} \;=\; \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_N x_N^2,$$

and $\det\left[\boxed{\mathrm{A}}\right] \;=\; \lambda_1 \cdot \lambda_2 \cdot \ldots \cdot \lambda_N$. Thus,

$$\begin{aligned} \rho(\mathbf{x}) &= \sqrt{\frac{\lambda_1 \cdot \ldots \cdot \lambda_N}{(2\pi)^N}} \exp\left[-\left(\lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_N x_N^2\right)\right] \\ &= \sqrt{\frac{\lambda_1}{2\pi}} \cdot \exp\left[-\lambda_1 \cdot x_1^2\right] \cdot \sqrt{\frac{\lambda_2}{2\pi}} \cdot \exp\left[-\lambda_2 \cdot x_2^2\right] \cdot \ldots \cdot \sqrt{\frac{\lambda_N}{2\pi}} \cdot \exp\left[-\lambda_N \cdot x_N^2\right] \\ &= \rho_1(x_1) \cdot \rho_2(x_2) \cdot \ldots \cdot \rho_N(x_N) \end{aligned}$$

where, for each n, $\rho_n : \mathbb{R} \longrightarrow \mathbb{R}$ is a univariate normal distribution with standard deviation $\sigma_n = \sqrt{\frac{1}{\lambda_n}}$. In the language of probability theory, we say that $x_1, \ldots, x_N$ are *independent* random variables, having normal distributions with standard deviations $\sigma_1, \ldots, \sigma_N$.

Given an arbitrary multivariate normal distribution $\rho(\mathbf{x})$, it would be nice if we could find some coordinate-system on $\mathbb{R}^N$ so that, with respect to this coordinate system, ρ was clearly the product of independent univariate distributions.

Proposition 35: *Principle Axis Theorem for Normal Distributions*
Let $\rho : \mathbb{R}^D \longrightarrow \mathbb{R}$ be a multivariate normal distribution with covariance matrix $\boxed{\mathsf{A}}$.

1. There is an orthogonal transformation $T : \mathbb{R}^D \longrightarrow \mathbb{R}^D$ so that, if $\eta : \mathbb{R}^D \longrightarrow \mathbb{R}$ is defined by $\eta(\mathbf{x}) = \rho(T(\mathbf{x}))$, then

$$\eta(x_1, \ldots, x_N) \;=\; \eta_1(x_1) \cdot \eta_2(x_2) \cdot \ldots \cdot \eta_D(x_D),$$

where $\eta_1, \ldots, \eta_D : \mathbb{R} \longrightarrow \mathbb{R}$ are univariate normal distributions with standard deviations $\sigma_1, \ldots, \sigma_D$, where $\sigma_d = \sqrt{\frac{1}{\lambda_d}}$, and $\lambda_1, \ldots, \lambda_D$ are the *eigenvalues* of $\boxed{\mathsf{A}}$.

In the language of probability theory, we can reformulate this result:

2. Let $\mathbf{X} \in \mathbb{R}^D$ be a *random vector* with distribution ρ. Then $\mathbf{X} = T(\mathbf{Y})$, where $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_D)$, and where $\mathbf{Y}_1, \ldots, \mathbf{Y}_D \in \mathbb{R}$ are *independent* normal random scalars, with standard deviations $\sigma_1, \ldots, \sigma_D$ respectively.

Proof: Exercise. Use the Principal Axes Theorem for quadratic forms (Theorem 5 on page 5).

□ [Proposition 35]