Symmetric matrices

We define a matrix $\vec{A}_{n\times n}$ as **symmetric** iff $\vec{A} = \vec{A}^T$, that is, $a_{ij} = a_{ji}\ \forall 1 \leq i, j \leq n$.

Symmetric matrices have various special properties. For example, it can be shown, given two distinct eigenvalues λ_1 and λ_2 and their corresponding eigenvectors $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$, that $\vec{v}_1 \perp \vec{v}_2$.

Proof. It is clear that $\lambda_1(\vec{v}_1 \cdot \vec{v}_2) = (\lambda_1\vec{v}_1) \cdot \vec{v}_2 = (\vec{A}\vec{v}_1) \cdot \vec{v}_2$. Using the definition of the dot product $\vec{a} \cdot \vec{b} = \vec{a}^T\vec{b}$, and the fact that $(\vec{P}\vec{Q})^T = \vec{Q}^T\vec{P}^T$, we see that $(\vec{A}\vec{v}_1)\cdot\vec{v}_2 = \vec{v}_1\cdot(\vec{A}^T\vec{v}_2) = \vec{v}_1\cdot(\vec{A}\vec{v}_2) = \vec{v}_1\cdot(\lambda_2\vec{v}_2) = \lambda_2(\vec{v}_1\cdot\vec{v}_2)$. Therefore, since $\lambda_1 \neq \lambda_2$ and $\lambda_1(\vec{v}_1 \cdot \vec{v}_2) = \lambda_2(\vec{v}_1 \cdot \vec{v}_2)$, we have $\vec{v}_1 \cdot \vec{v}_2 = 0 \implies \vec{v}_1 \perp \vec{v}_2$. $\square$

We can also show that symmetric matrices are orthogonally diagonalisable, that is, they are diagonalisable in the form $\vec{A} = \vec{P}\vec{D}\vec{P}$ where $\vec{P}$ is an orthogonal matrix, i.e. $\vec{P}\vec{P}^T = \vec{I}_n$. The proof of this fact is, however, beyond the scope of this course. A very important fact involved in the proof is the **spectral theorem** for symmetric matrices, which states that, counting multiplicities, a symmetric matrix $\vec{A}_{n\times n}$ has n distinct **real** eigenvalues, and that for any eigenvalue λ, its algebraic and geometric multiplicities are equal. It also states a result that we already proved: that any two distinct eigenspaces of $\vec{A}$ are orthogonal.

A **symmetric spectral decomposition** of $\vec{A}_{n\times n}$ is $\vec{A} = \sum_{i=1}^{n} \lambda_i(\vec{u}_i\vec{u}_i^T)$, where $\vec{u}_i$ is the ith column of the $\vec{P}$ matrix in the orthogonal diagonalisation of $\vec{A}$.

Quadratic forms

A quadratic form is a function Q of the form

$$\begin{cases} \mathbb{R}^n \to \mathbb{R} \\ \vec{x} \mapsto \vec{x}^T\vec{A}\vec{x} \end{cases}$$

For example, given

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \vec{A}_1 = \begin{bmatrix} 7 & 0 \\ 0 & 4 \end{bmatrix} \vec{A}_2 = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$$

we would have

$$Q_1(\vec{x}) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 7 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 7x_1^2 + 4x_2^2$$

and

$$Q_2(\vec{x}) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 5x_2^2 + 6x_1x_2$$

Denoting $\vec{A} = (a_{ij})$, $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$, we can remark that

$$Q(\vec{x}) = \vec{x}^T\vec{A}\vec{x} = \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{i<j}(a_{ij} + a_{ji})x_ix_j$$

noting that, if $\vec{A}$ is symmetric, then $\sum_{i<j}(a_{ij} + a_{ji})x_ix_j = 0$. Also, consider the function

$$\langle \cdot, \cdot \rangle \begin{cases} \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \\ (\vec{x}, \vec{y}) \mapsto \vec{x}^T\vec{A}\vec{y} \end{cases}$$

This is an inner product if and only if the four properties of inner products hold. We can show that $\vec{A} = \vec{A}^T \iff \vec{x}^T\vec{A}\vec{y} = \vec{y}^T\vec{A}\vec{x}$, that is, if $\vec{A}$ is symmetric. This directly implies that the function is distributive and linear with regards to multiplication. Finally, since we know that $\vec{A} = \vec{A}^T \iff \sum_{i<j}(a_{ij} + a_{ji})x_ix_j = 0$, it follows that $\langle \vec{x}, \vec{x} \rangle = \vec{x}^T\vec{A}\vec{x} \geq 0$.

Note: if we use a basis for $\mathbb{R}^n B = \{\vec{u}_1, \vec{u}_2 \cdots \vec{u}_n\}$, instead of the usual orthonormal basis, then in the new basis out coordinates for $\vec{x} \in mathbbR^n$ are $\vec{y} = [\vec{x}]_B = \vec{P}^{-1}\vec{x}$, since we know that $\vec{x} = \vec{P}\vec{y}$. Then, we have:

$$\vec{x}^T\vec{A}\vec{x} = (\vec{P}\vec{y})^T\vec{A}(\vec{P}\vec{y}) = \vec{y}^T(\vec{P}^T\vec{A}\vec{P})\vec{y} = \vec{y}^T\vec{D}\vec{y}$$

And so, we have the following fact: if $B = \{\vec{u}_1, \vec{u}_2 \cdots \vec{u}_n\}$ is an orthonogmal eigenbasis for $\mathbb{R}^n$, i.e. comprised of the columns of $\vec{P}$ from the orthogonal diagonalisation, then $Q(\vec{x}) = \vec{x}^T\vec{A}\vec{x} = \vec{y}^T\vec{D}\vec{y}$, where $\vec{D}$ is a diagonal matrix with the corresponding eigenvalues on its diagonal. This allows us to use a change of variable to convert any such quadratic form Q into another quadratic form with a symmetric matrix, which eliminates the third term in the formula.

For any such Q, we say that it (or its matrix $\vec{A}$) is either:

1. **Positive definite** if $Q(\vec{x}) \geq 0 \; \forall \; \vec{x} \in \mathbb{R}^n$

2. **Negative definite** if $Q(\vec{x}) \leq 0 \ \forall \ \vec{x} \in \mathbb{R}^n$
3. **Indefinite** if $\exists \vec{x} \in \mathbb{R}^n : Q(\vec{x}) > 0$ and $\exists \vec{y} \in \mathbb{R}^n : Q(\vec{y}) < 0$

So, we can summarise that $\langle \vec{x}, \vec{y} \rangle = \vec{x}^T \vec{A} \vec{y}$ is indeed an inner product iff $Q(\vec{x})$ is positive definite.

We can also say that $\vec{A}$ is positive definite iff all its eigenvalues are positive, negative definite iff they are all negative, and indefinite if neither is the case.

Proof. By choosing $\vec{y} = \vec{P}^T \vec{x}, \vec{P}^T = \vec{P}^{-1}$ and defining $\vec{D}$ as a diagonal matrix with the eigenvalues of $\vec{A}$ on its diagonal, we have $Q(\vec{x}) = \vec{y}^T \vec{D} \vec{y} = \sum_{i=1}^{n} \lambda_i y_i^2$. Also, since $\vec{P}$ is invertible and $\vec{y} = \vec{P}^T \vec{x}$, we know that $\vec{x} = 0 \iff \vec{y} = 0$ and Q is both injective and surjective. From these two facts, we see that $Q(\vec{x}) \geq 0 \ \forall \ \vec{x} \in \mathbb{R}^n$ iff all λ_i are positive, and $Q(\vec{x}) \leq 0 \ \forall \ \vec{x} \in \mathbb{R}^n$ iff all λ_i are negative. If some eigenvalues are positive and some are negative, then $Q(\vec{x})$ can be positive or negative, and so Q is indefinite. $\square$

Note: if we use a basis for $\mathbb{R}^n B = \{\vec{u}_1, \vec{u}_2 \cdots \vec{u}_n\}$, instead of the usual orthonormal basis, then in the new basis our coordinates for $\vec{x} \in mathbbR^n$ are $\vec{y} = [\vec{x}]_B = \vec{P}^{-1} \vec{x}$, since we know that $\vec{x} = \vec{P} \vec{y}$. Then, we have: