Symmetric Matrices

Ziv Yaniv

Version 1.0
March 20 2012

Theorem 0.1 *The eigenvalues of a real symmetric matrix are real.*

Proof Given the symmetric real matrix A we have:

$$A\mathbf{x} = \lambda\mathbf{x} \tag{1}$$

Where λ is an eigenvalue and $\mathbf{x}$ the corresponding eigenvector. Until we prove the theorem we must assume that λ might be a complex number ($\lambda = a+ib$) and $\mathbf{x}$ might contain components which are complex too. Remembering that $\overline{\lambda\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$ and that $A = \overline{A} = A^T$ we take the conjugates of equation 1:
$\overline{A\mathbf{x}} = \overline{\lambda\mathbf{x}}$ leads to $A\overline{\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$. Transpose this to get $\overline{\mathbf{x}}^T A = \overline{\mathbf{x}}^T\overline{\lambda}$ Now taking the dot product of the first equation with $\overline{\mathbf{x}}$ and the last equation with $\mathbf{x}$ we get:
$\overline{\mathbf{x}}^T A\mathbf{x} = \overline{\mathbf{x}}^T\lambda\mathbf{x}$ and $\overline{\mathbf{x}}^T A\mathbf{x} = \overline{\mathbf{x}}^T\overline{\lambda}\mathbf{x}$
Which gives us: $\overline{\mathbf{x}}^T\lambda\mathbf{x} = \overline{\mathbf{x}}^T\overline{\lambda}\mathbf{x}$
Therefore $\lambda = \overline{\lambda}$ proving that λ is real ($a + bi = a - bi$, so the complex coefficient is equal to zero).

Theorem 0.2 *The eigenvectors of a real symmetric matrix which correspond to different* [1] *eigenvalues are perpendicular.*

Proof Let λ_1 and λ_2 be two different eigenvalues and $\mathbf{x_1}$ and $\mathbf{x_2}$ the corresponding eigenvectors. This gives us the following two equations:

$$A\mathbf{x_1} = \lambda_1\mathbf{x_1}$$
$$A\mathbf{x_2} = \lambda_2\mathbf{x_2}$$

Taking the dot product with $\mathbf{x_2}$ we get:

$$(\lambda_1\mathbf{x_1})^T\mathbf{x_2} = (A\mathbf{x_1})^T\mathbf{x_2} = \mathbf{x_1}^T A^T\mathbf{x_2} = \mathbf{x_1}^T A\mathbf{x_2} = \mathbf{x_1}^T\lambda_2\mathbf{x_2}$$

The left side is $\mathbf{x_1}^T\lambda_1\mathbf{x_2}$ and the right side is $\mathbf{x_1}^T\lambda_2\mathbf{x_2}$. Since $\lambda_1 \neq \lambda_2$ this proves that $\mathbf{x_1}^T\mathbf{x_2} = 0$. The eigenvectors are perpendicular.

[1] Proof that *all eigenvectors of a real symmetric matrix are orthogonal to each other can be found in [1]*

Theorem 0.3 *Given a real symmetric matrix A the solution to the optimization problem:*

1. $$\min_{\mathbf{x}} \mathbf{x}^T A\mathbf{x}, \quad s.t. \|\mathbf{x}\| = 1$$

   is the eigenvector corresponding to the minimal eigenvalue.

2. $$\max_{\mathbf{x}} \mathbf{x}^T A\mathbf{x}, \quad s.t. \|\mathbf{x}\| = 1$$

   is the eigenvector corresponding to the maximal eigenvalue.

Proof 1. Using Lagrange multipliers the optimization problem is rewritten as:

$$\min_{\mathbf{x}} L(\mathbf{x}, \lambda) = \mathbf{x}^T A\mathbf{x} + \lambda(1 - \mathbf{x}^T\mathbf{x})$$

where at the optimum we have the following necessary conditions:

$$\begin{aligned} \tfrac{\partial L}{\partial \mathbf{x}} &= (A\mathbf{x} + \mathbf{x}^T A) - 2\lambda\mathbf{x} = 0 \\ \tfrac{\partial L}{\partial \lambda} &= 1 - \mathbf{x}^T\mathbf{x} = 0 \end{aligned} \tag{2}$$

as A is symmetric we have:

$$2A\mathbf{x} = 2\lambda\mathbf{x}$$

The minimum is thus obtained for $\mathbf{x}$ which is an eigenvector of the matrix A. There are n mutually orthogonal (see Theorem 0.2) eigenvectors $\mathbf{e_i}$ associated with A, that span $\mathbb{R}^n$. Which means that $\forall\mathbf{x},\ \mathbf{x} = \sum w_i\mathbf{e_i}$.

Looking back at our original problem we have:

$$\mathbf{x}^T A\mathbf{x} = \left(\sum w_i\mathbf{e_i}^T\right) A \left(\sum w_i\mathbf{e_i}\right) = \sum w_i\mathbf{e_i}^T \sum w_i\lambda_i\mathbf{e_i} = \sum \lambda_i w_i^2$$

Arranging the eigenvectors in order $\lambda_1 \geq \ldots \geq \lambda_n$ (possible as they are real values, see Theorem 0.1) we have:

$$\sum \lambda_n w_i^2 \leq \sum \lambda_i w_i^2 = \mathbf{x}^T A\mathbf{x}$$

At the minimum we also have:

$$\mathbf{x}^T\mathbf{x} = \sum w_i^2 = 1$$

Plugging this back into the previous equation we get:

$$\lambda_n = \sum \lambda_n w_i^2 \leq \mathbf{x}^T A\mathbf{x}$$

The minimal value is bounded from below by the smallest eigenvalue and is attained when $\mathbf{x} = \mathbf{e_n}$.

2. The same proof, only now the quadratic term, $\mathbf{x}^T A\mathbf{x}$, is bounded from above.

Theorem 0.4 *Let A be a real nxn symmetric matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and corresponding orthonormal eigenvectors $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}$ (see Theorem 0.2). Define $X_k = (\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_k})$ $(k = 1, 2, \ldots, n-1)$ and $X = (\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n})$. Then if we assume that $\alpha \neq 0$, we have the following:*

1.
$$\sup_{\alpha} \left\{ \frac{\alpha^T A \alpha}{\alpha^T \alpha} \right\} = \lambda_1$$
and the supremum is attained if $\alpha = \mathbf{x_1}$.

2.
$$\sup_{X_k^T \alpha = 0} \left\{ \frac{\alpha^T A \alpha}{\alpha^T \alpha} \right\} = \lambda_{k+1}$$
and the supreumum is attained if $\alpha = \mathbf{x_{k+1}}$.

3.
$$\inf_{\alpha} \left\{ \frac{\alpha^T A \alpha}{\alpha^T \alpha} \right\} = \lambda_n$$
and the infimum is attained if $\alpha = \mathbf{x_n}$.

4. *If $X_{n-k} = (\mathbf{x_{n-k+1}}, \mathbf{x_{n-k+2}}, \ldots, \mathbf{x_n})$ then*
$$\inf_{X_{n-k}^T \alpha = 0} \left\{ \frac{\alpha^T A \alpha}{\alpha^T \alpha} \right\} = \lambda_{n-k}$$
and the infimum is attained if $\alpha = \mathbf{x_{n-k}}$.

Proof 1. Let $\alpha = X\mathbf{y} = y_1\mathbf{x_1} + y_2\mathbf{x_2} + \cdots + y_n\mathbf{x_n}$ and
$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$
Then
$$\frac{\alpha^T A \alpha}{\alpha^T \alpha} = \frac{\mathbf{y}^T X^T A X \mathbf{y}}{\mathbf{y}^T \mathbf{y}} = \frac{\mathbf{y}^T X^T X \Lambda \mathbf{y}}{\mathbf{y}^T \mathbf{y}} = \frac{(\sum_i \lambda_i \mathbf{y_i}^2)}{\mathbf{y}^T \mathbf{y}} \leq \frac{\lambda_1 \mathbf{y}^T \mathbf{y}}{\mathbf{y}^T \mathbf{y}} = \lambda_1$$
with equality when $y_1 = 1, y_2 = y_3 = \cdots = y_n = 0$, that is, when $\alpha = \mathbf{x_1}$.

2. If $\alpha \perp \mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_k}$, then $y_1 = y_2 = \cdots = y_k = 0$. The result then follows with the same argument as 1.

3. Same proof as 1 but with the inequality reversed.

4. Same proof as 2 but with inequality reversed.

1 Matrices of the form AA^T

Theorem 1.1 *Matrices of the form AA^T, where A is non-singular, have the following properties:*

1. *They are positive definite.*
2. *They have positive eigenvalues.*
3. *The Singular Value Decomposition (SVD) of A yields the eigenvalues and eigenvectors of AA^T.*

Proof

1. A matrix B is positive definite if:

$$\forall \mathbf{x},\ \mathbf{x} \neq 0 \quad \mathbf{x}^T B \mathbf{x} > 0$$

Given a matrix B of the form AA^T we have:

$$\mathbf{x}^T B \mathbf{x} = \mathbf{x}^T A A^T \mathbf{x} = (A^T \mathbf{x})^T (A^T \mathbf{x}) = \|A^T \mathbf{x}\|^2 > 0$$

2. Given a matrix B of the form AA^T with eigenvalue λ and corresponding eigenvector $\mathbf{x}$ we have:

$$B\mathbf{x} = \lambda \mathbf{x}$$

Premultiplying by $\mathbf{x}^T$ we get:

$$\mathbf{x}^T B \mathbf{x} = \lambda \mathbf{x}^T \mathbf{x} = \lambda \|\mathbf{x}\|^2$$

As B is positive definite we have:

$$\begin{array}{rcl} \lambda\|\mathbf{x}\|^2 & > & 0 \\ & \Downarrow & \\ \lambda & > & 0 \end{array}$$

3. The Singular Value Decomposition of the matrix A is given by

$$A_{m\times n} = U_{m\times n} S_{n\times n} V^T_{n\times n}$$

where the columns of U are an orthonormal basis for the column space of A, the rows of V are an orthonormal basis for the row space of A and S is a diagonal matrix. We now look at the matrix AA^T:

$$AA^T = (USV^T)(VSU^T) = US^2U^T$$

Postmultipying this equation by U yields the following equation:

$$(AA^T)U = US^2$$

The eigenvectors of AA^T are the columns of U with corrosponding eigenvalues in S.

2 Eigenvalues/Eigenvectors of a 2×2 Symmetric Matrix

Given a symmetric matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}$$

We obtain the eigenvalues of A by solving the characteristic equation:

$$det(A - \lambda I) = 0$$

For the matrix A this is a quadratic equation:

$$(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}^2 = \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}^2)$$

whose solution yields the eigenvalues:

$$\lambda_1 = \frac{a_{11} + a_{22} + \sqrt{(a_{11} - a_{22})^2 + 4a_{12}^2}}{2} \qquad \lambda_2 = \frac{a_{11} + a_{22} - \sqrt{(a_{11} - a_{22})^2 + 4a_{12}^2}}{2}$$

and corresponding eigenvectors:

$$\mathbf{v_1} = [\lambda_1 - a_{22}, a12] \qquad \mathbf{v_2} = [-a12, \lambda_1 - a_{22}]$$

Note that when $a_{12} \to 0$ the eigenvectors are:

$$\begin{array}{cc} a_{11} \to 0 & a_{22} \to 0 \\ \mathbf{v_1} = [0,1]\ \mathbf{v_2} = [1,0] & \mathbf{v_1} = [1,0]\ \mathbf{v_2} = [0,1] \end{array}$$

References

[1] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, third edition, 1996.