Diagonalization of Symmetric Matrices

Theorem: Let A be an $n \times n$ matrix with real entries. Then the following conditions are equivalent.

1. (a) A is symmetric: $A^T = A$.
   (b) For any $X, Y \in \mathbf{R}^n$, $(AX|Y) = (X|AY)$.
   (c) The matrix of T_A with respect to *every* orthonormal basis is symmetric.

2. There exists an orthogonal basis of $\mathbf{R}^n$ consisting of eigenvectors of A.

3. There exist an invertible matrix S such that $S^{-1} = S^T$ and a (real) diagonal matrix D such that $A = SDS^{-1}$.

Hermitian inner product

If $B \in M_{mn}(\mathbf{C})$, $B^* := \overline{B}^T \in M_{nm}$.
If $X, Y, \in \mathbf{C}^n$, $(X|Y) := \sum X_i \overline{Y}_i = Y^* X = \overline{X^* Y}$.
This is similar to the formula for the inner product in the real case, except for the complex conjugate. This is necessary for the positivity condition.

Properties of the inner product

1. $(X + X'|Y) = (X|Y) + (X'|Y)$ and
   $(X|Y + Y') = (X|Y) + (X|Y')$.

2. $(\lambda X|Y) = \lambda(X|Y) + (X|\overline{\lambda}Y)$.

3. $(X|Y) = \overline{(Y|X)}$.

4. $(X|X) > 0$ if $X \neq \mathbf{0}$.

Formula: If $X \in \mathbf{C}^n, Y \in \mathbf{C}^m$, $(BX|Y) = (X|B^* Y)$.

The following result reduces to the theorem above when the entries of A are real.

Variant: Let A be an $n \times n$ matrix with complex entries. Then the following conditions are equivalent.

1. (a) A is Hermitian: $A^* = A$.
   (b) For any $X, Y \in \mathbf{C}$, $(AX|Y) = (X|AY)$
   (c) The matrix of T_A with respect to *every* orthonormal basis for $\mathbf{C}^n$ is hermitian.

2. There exists an orthogonal basis of $\mathbf{C}^n$ consisting of eigenvectors of A, and the eigenvalue of A are real.

3. There exist an invertible matrix S such that $S^{-1} = S^*$ and a real diagonal matrix D such that $A = SDS^{-1}$.

Proofs

First we prove the "formula."

$$(X|B^*Y) := (B^*Y)^*X = (Y^*B^{**})X =$$

$$(Y^*B)X = Y^*(BX) = (BX|Y)$$

From now on, let A be an $n \times n$ matrix with complex entries. Let us prove the equivalence of the three conditions in (1). If (a) holds, then $(AX|Y) = (X|A^*Y) = (X|AY)$, so (b) holds. If $\mathcal{B} := \mathbf{v}_1, \cdots \mathbf{v}_n$ is any basis for $\mathbf{C}^n$, then the matrix for T_A with respect to $\mathcal{B}$ is the matrix A'_{ij} such that $A\mathbf{v}_j = \sum A'_{ij}\mathbf{v}_i$. If $\mathcal{B}$ is orthonormal and (b) holds, then

$$A'_{ij} = (A\mathbf{v}_j|\mathbf{v}_i) = (\mathbf{v}_j|A\mathbf{v}_i) = \overline{(\mathbf{v}_i|A\mathbf{v}_j)} = \overline{A}'_{ji}.$$

This says that $A' = A'^*$. Evidently (c) is a special case of (a).

Now let's prove that (2) implies (3). Let $\mathcal{B} := (\mathbf{v}_1, \ldots \mathbf{v}_n)$ be an othogonal basis of $\mathbf{C}^n$ consisting of eigenvectors of A. For each i, we may replace $\mathbf{v}_i$ by $\mathbf{v}_i/||\mathbf{v}_i||$, so that now $\mathcal{B}$ is an orthonormal basis. Since $\mathbf{v}_i$ is still an eigenvector of A, $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for some λ_i, which by hypothesis is in fact real. Let S be the matrix whose jth column is $\mathbf{v}_j$. We claim that $S^{-1} = S^*$, or equivalently, that $S^*S = I$. But the ijth entry of S^*S is

$$R_i(S^*)C_j(S) = C_i(\overline{S})^T C_j(S) = C_i(S)^* C_j(S) = (\mathbf{v}_i|\mathbf{v}_j) = \delta_{ij}.$$

It follows that S has the desired proporties.

The fact that (3) implies (1) can be deduced from the equivalences of the three parts of (1), but let's prove it directly. If $A = SDS^{-1}$ where $S^{-1} = S^*$ and D is diagonal and real, then $D = D^*$, and so

$$A^* = (SDS^*)^* = S^{**}D^*S^* = SDS^* = A.$$

The proof that (1) implies (2) is more difficult.

Lemma: If A is Hermitian, then all the eigenvalues of A are real.

Proof: If λ is an eigenvalues of A, then there is a nonzero $\mathbf{v}$ such that $A\mathbf{v} = \lambda\mathbf{v}$. Hence

$$\lambda(\mathbf{v}|\mathbf{v}) = (\lambda\mathbf{v}|\mathbf{v}) = (A\mathbf{v}|\mathbf{v}) = (\mathbf{v}|A\mathbf{v}) = (\mathbf{v}|\lambda\mathbf{v}) = \overline{\lambda}(\mathbf{v}|\mathbf{v}).$$

Since $(\mathbf{v}|\mathbf{v}) \neq 0$, it follows that $\lambda = \overline{\lambda}$, so λ is real.

Although the following result is really a consequence of the theorem, it is worth proving seperately.

Lemma: If A is Hermitian and $\mathbf{v}$ and $\mathbf{v}'$ are eigenvectors of A corresponding to distinct eigenvalues, then $\mathbf{v}$ and $\mathbf{v}'$ are orthogonal.

Proof: Say $A\mathbf{v} = \lambda\mathbf{v}$ and $A\mathbf{v}' = \lambda'\mathbf{v}'$. Recall that λ and λ' are real. Then

$$\lambda(\mathbf{v}|\mathbf{v}') = (\lambda\mathbf{v}|\mathbf{v}') = (A\mathbf{v}|\mathbf{v}') = (\mathbf{v}|A\mathbf{v}') = (\mathbf{v}|\lambda'\mathbf{v}') = \lambda'(\mathbf{v}|\mathbf{v}')$$

Thus $(\lambda - \lambda')(\mathbf{v}|\mathbf{v}') = 0$. If $\lambda \neq \lambda'$, this implies that $(\mathbf{v}|\mathbf{v}') = 0$.

Lemma: If A is Hermitian and $\mathbf{v}$ is an eigenvector of A, let $W := \mathbf{v}^{\perp}$. Then if $\mathbf{w} \in W$, $A\mathbf{w} \in W$.

Proof: If $\mathbf{w} \in W$, then

$$(A\mathbf{w}|\mathbf{v}) = (\mathbf{w}|A\mathbf{v}) = (\mathbf{w}|\lambda\mathbf{v}) = \overline{\lambda}(\mathbf{w}|\mathbf{v}) = 0.$$

Hence $A\mathbf{w} \perp \mathbf{v}$, that is $A\mathbf{w} \in w$.

Now we prove that (1) implies (2) by induction on n. If $n = 1$ there is nothing to prove. Say $n > 1$. Let λ be a root of $p_A(X)$. Then there exists a $\mathbf{v}$ such that $A\mathbf{v} = \lambda\mathbf{v}$ and $||\mathbf{v}|| = 1$. Let $W := \mathbf{v}^{\perp}$. This is a vector space of dimension $n - 1$. We can choose an orthonormal basis $\mathcal{B} := (\mathbf{v}_1, \ldots \mathbf{v}_n)$ for $\mathbf{C}^n$ with $\mathbf{v}_1 = \mathbf{v}$. Then $(\mathbf{v}_2, \ldots \mathbf{v}_n)$ is an orthnormal basis for W. By (1c), the matrix A' for T_A with respect to $\mathcal{B}$ is still Hermitian. Since $\mathbf{v}_1$ is an eigenvector for A, $A'_{i1} = 0$ if $j \neq 1$. By (2), $A'_{1j} = 0$ if $j \neq 1$. The induction hypothesis says that W has an orthogonal basis consisting of eigenvectors of the remaining part of the matrix, and this concludes the proof.