Lecture 15. Diagonalization of Quadratic Forms

December 2022

Suppose $f : V \to V$ is a linear operator on a finite-dimensional inner product space V having the adjoint $f^* : V \to V$.

LEMMA 15.1 Any λ-eigenvector of f is orthogonal to any μ-eigenvector of f^* whenever $\lambda \neq \mu$.

Proof. Suppose $f\mathbf{x} = \lambda\mathbf{x}$ and $f^*\mathbf{y} = \mu\mathbf{y}$ and $\lambda \neq \mu$. Then

$$\lambda\langle \mathbf{x}, \mathbf{y}\rangle = \langle f\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, f^*\mathbf{y}\rangle = \mu\langle \mathbf{x}, \mathbf{y}\rangle.$$

As $\lambda \neq \mu$, the last equality is only possible if $\langle \mathbf{x}, \mathbf{y}\rangle = 0$. □

LEMMA 15.2 If $U \subset V$ is f-invariant, then $U^\perp$ is f^*-invariant.

Proof. By definition of an invariant subspace, $\mathbf{x} \in U$ implies $f\mathbf{x} \in U$. Take any $\mathbf{y} \in U^\perp$. Since for any $\mathbf{x} \in U$ we have

$$\langle \mathbf{x}, f^*\mathbf{y}\rangle = \langle f\mathbf{x}, \mathbf{y}\rangle = 0,$$

then $f^*\mathbf{y} \in U^\perp$. □

THEOREM 15.3 (Spectral Theorem) Let $f : V \to V$ be a linear operator on a finite-dimensional inner product space V.

1. If f is self-adjoint, then f is orthogonally diagonalizable, i.e., there exists an orthonormal basis of V in which the matrix of f is diagonal.
2. Conversely, if f has a diagonal matrix with respect to some orthonormal basis, then f is self-adjoint.

Proof. 2 $\implies$ 1 follows from Proposition 13.11.

1 $\implies$ 2 If f is self-adjoint, then f has an $n \times n$ symmetric matrix A with respect to some orthonormal basis.

We proceed by induction on n. If $n = 1$, A is already diagonal.

If $n > 1$, assume that **1 $\implies$ 2** for $(n-1) \times (n-1)$ symmetric matrices. By Theorem 5.23, A has a (real) eigenvalue λ_1, and let $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1$, where $\|\mathbf{x}_1\| = 1$. Use the Gram-Schmidt algorithm to find an orthonormal basis $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ for $\mathbb{R}^n$. The subspace $U = \operatorname{span}\{\mathbf{x}_1\}$ is f-invariant, so by Lemma 15.2, $U^\perp = \operatorname{span}\{\mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is f-invariant because $f^* = f$.

Therefore, V is represented as a direct sum of invariant subspaces, $V = U \oplus U^\perp$, and the matrix of f with respect to the basis $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is $\begin{pmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{pmatrix}$ in block form. By induction, in $U^\perp = \operatorname{span}\{\mathbf{x}_2, \ldots, \mathbf{x}_n\}$ there exists an orthonormal basis of eigenvectors of $f|_{U^\perp}$. Thus, f has an orthonormal basis of n eigenvectors. □

The proof of the following theorem will be omitted.

THEOREM 15.4 (Spectral Theorem in Complex Vector Spaces) If A is Hermitian, there is a unitary matrix U such that $U^H AU$ is diagonal.

Diagonalization of Quadratic Forms

THEOREM 15.5 Let $b(\mathbf{x}, \mathbf{y})$ be a bilinear form on an inner product space V. Then there exists a unique linear operator $f : V \to V$ such that

$$b(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, f\mathbf{y} \rangle$$

for all $\mathbf{x}, \mathbf{y} \in V$.

Proof. Since for a fixed $\mathbf{y}$ the function $b(\mathbf{x}, \mathbf{y})$ is linear with respect to $\mathbf{x}$, then by Riesz representation theorem, there exists a unique vector $f\mathbf{y}$ such that $b(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, f\mathbf{y} \rangle$. It is easy to prove that f is linear. □

DEFINITION 15.6 A linear operator $f : V \to V$ is said to be **associated** with a bilinear form b if

$$b(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, f\mathbf{y} \rangle$$

for all $\mathbf{x}, \mathbf{y} \in V$.

THEOREM 15.7 The matrices of a bilinear form b and an associated linear operator $f : V \to V$ are related by $B = GA$, where G is a Gram matrix.

Proof. We have $b(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T B\mathbf{y}$ and $\langle \mathbf{x}, f\mathbf{y} \rangle = \mathbf{x}^T GA\mathbf{y}$. As

$$\mathbf{x}^T B\mathbf{y} = \mathbf{x}^T GA\mathbf{y}$$

for any two vectors $\mathbf{x}$ and $\mathbf{y}$, we have $B = GA$. □

COROLLARY 15.8 If the basis in V is orthonormal, then $A = B$.

We shall assume from now on that any quadratic form in the variables $x_1, x_2, \ldots, x_n$ is given by

$$q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x},$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ and A is a symmetric $n \times n$ matrix. Given such a form, the problem is to find an orthonormal basis with the property that when q is expressed in terms of new variables, there are no cross terms. If we write $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T$, this amounts to asking that $q = \mathbf{y}^T D\mathbf{y}$ where D is diagonal. It turns out that this can always be accomplished. Actually, the following theorem is valid.

THEOREM 15.9 (Principal Axis Theorem) Let V be an inner product space. Then for any quadratic form, there is an orthonormal basis, relative to which it is described by a diagonal matrix.

Proof. Consider any orthonormal basis. Suppose $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, where A is a symmetric matrix. Consider a bilinear form $b(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A\mathbf{y}$ and the associated linear operator $f : V \to V$ described by the same matrix. Since the matrix A is symmetric, it can be orthogonally diagonalized. In fact a matrix P can be found that is orthogonal (i.e., $P^{-1} = P^T$) and diagonalizes A:

$$P^T AP = D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n). \quad \square$$

DEFINITION 15.10 Let $q = \mathbf{x}^T A\mathbf{x}$ be a quadratic form where A is a symmetric matrix and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the (real) eigenvalues of A repeated according to their multiplicities. A corresponding set $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ of orthonormal eigenvectors for A is called a set of **principal axes** for the quadratic form q.

EXAMPLE 15.11 Find an orthonormal basis, in which the quadratic form given by

$$q = x_1^2 + 4x_1x_2 + x_2^2$$

has a diagonal form. Find the corresponding principal axes.

Solution. The form can be written as $q = \mathbf{x}^T A\mathbf{x}$, where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

A routine calculation yields $p_A(\lambda) = \det(A - \lambda I) = (\lambda - 3)(\lambda + 1)$, so the eigenvalues are $\lambda_1 = 3$, $\lambda_2 = -1$. Corresponding orthonormal eigenvectors are the principal axes:

$$\mathbf{f}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \mathbf{f}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

The matrix

$$P = (\mathbf{f}_1 \quad \mathbf{f}_2) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

is an orthogonal transition matrix. In terms of new variables, y_1 and y_2, the quadratic form is diagonal:

$$q = 3y_1^2 - y_2^2.$$

THEOREM 15.12 (Simultaneous Diagonalization Theorem) Let h and g be two quadratic forms in a finite-dimensional vector space V, the form g being positive definite. Then there is a basis, relative to which both forms are described by diagonal matrices.

Proof. Introduce an inner product in V by letting

$$\langle \mathbf{x}, \mathbf{y} \rangle = g(\mathbf{x}, \mathbf{y}).$$

By Theorem 15.9, there exists an orthonormal basis, relative to which h is described by a diagonal matrix. □

Algorithm for simultaneous diagonalization of two quadratic forms.

1. Suppose quadratic forms h and g are described by matrices H and G, respectively. Since g is positive definite, the matrix G can be considered as Gram matrix. Let A be the matrix of a linear operator associated with h. Then it is given by $A = G^{-1}H$ by Theorem 15.7.
2. To find eigenvalues and eigenvectors of A, we must solve the characteristic equation $\det(A - \lambda I) = 0$. We have:

   $$\det(A - \lambda I) = \det(G^{-1}H - \lambda I) = \det(G^{-1}(H - \lambda G)) = \det(G^{-1})\det(H - \lambda G).$$

   Thus, the characteristic equation has the same roots as $\det(H - \lambda G) = 0$, which is called the **generalized characteristic equation**.

3. To find an eigenvector $\mathbf{x}_k$ associated with eigenvalue λ_k, we must solve vector equation $(A - \lambda_k I)\mathbf{x}_k = \mathbf{0}$, which is equivalent to $(H - \lambda_k G)\mathbf{x}_k = \mathbf{0}$.
4. Eigenvectors associated to different eigenvalues are orthogonal by Theorem 15.1. To find orthonormal bases for each eigenspace, we use Gram matrix G (the Gram–Schmidt algorithm may be needed).
5. Then the set of all these basis vectors is orthonormal with respect to g and contains n vectors.

EXAMPLE 15.13 Diagonalize simultaneously two quadratic forms:

$$h = 13x_1^2 - 10x_1x_2 + 3x_2^2, \qquad g = 11x_1^2 - 6x_1x_2 + x_2^2.$$

Solution. The matrices of h and g are

$$H = \begin{pmatrix} 13 & -5 \\ -5 & 3 \end{pmatrix}, \qquad G = \begin{pmatrix} 11 & -3 \\ -3 & 1 \end{pmatrix}.$$

Sylvester's criterion shows that G is positive definite. The generalized characteristic equation

$$\begin{vmatrix} 13 - 11\lambda & -5 + 3\lambda \\ -5 + 3\lambda & 3 - \lambda \end{vmatrix} = 2\lambda^2 - 16\lambda + 14 = 0$$

yields $\lambda_1 = 1$, $\lambda_2 = 7$. Corresponding eigenvectors are

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \mathbf{x}_2 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}.$$

Vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are orthogonal with respect to Gram matrix G. To normalize the vectors, find their norms using G:

$$|\mathbf{x}_1|^2 = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 11 & -3 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 6, \quad |\mathbf{x}_2|^2 = \begin{pmatrix} 1 & 4 \end{pmatrix} \begin{pmatrix} 11 & -3 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \end{pmatrix} = 3.$$

Now, the vectors

$$\mathbf{f}_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \mathbf{f}_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

form orthonormal basis, so the matrix

$$P = (\mathbf{f}_1 \quad \mathbf{f}_2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{3} & 4/\sqrt{3} \end{pmatrix}$$

is a transition matrix. In terms of new variables, y_1 and y_2, the quadratic forms are diagonal:

$$g = y_1^2 + y_2^2, \qquad h = y_1^2 + 7y_2^2.$$

EXAMPLE 15.14 If neither h nor g is positive definite, then we can't guarantee that they can be diagonalized simultaneously. For example, this doesn't hold for $h = x_1^2 - x_2^2$, $g = 2x_1x_2$.