Handout 3: Bilinear and Quadratic Forms

This handout should be read just before Chapter 4 of the textbook.

Endomorphisms of a Vector Space

This handout discusses some important constructions from linear algebra that we will use throughout the rest of the course. Most of the material discussed here is described in more detail in *Introduction to Linear Algebra* by Johnson, Riess and Arnold.

Let us begin by recalling a few basic ideas from linear algebra. Throughout this handout, let V be an n-dimensional vector space over the real numbers. (In all of our applications, V will be a linear subspace of some Euclidean space $\mathbb{R}^k$, that is, a subset that is closed under vector addition and scalar multiplication.)

A linear map from V to itself is called an ***endomorphism*** of V. Given an endomorphism $A\colon V \to V$ and a basis $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ for V, we can express the image under A of each basis vector as a linear combination of basis vectors:

$$A\mathbf{x}_j = \sum_{i=1}^{n} A_{ij}\mathbf{x}_i.$$

This determines an $n \times n$ matrix $A_{\mathbf{x}} = (A_{ij})$, called the ***matrix of A with respect to the given basis***. (Thanks to Eli Fender for suggesting this notation.) By linearity, the action of A on any other vector $v = \sum_j v_j\mathbf{x}_j$ is then determined by

$$A\left(\sum_{j=1}^{n} v_j\mathbf{x}_j\right) = \sum_{i,j=1}^{n} A_{ij}v_j\mathbf{x}_i.$$

If we associate with each vector $v = \sum_j v_j\mathbf{x}_j$ its n-tuple of coefficients $(v_1, \dots, v_n)$ arranged as a column matrix, then the n-tuple associated with $w = Av$ is determined by matrix multiplication:

$$\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Just as in the case of linear maps on $\mathbb{R}^n$, the jth column of this matrix is the n-tuple associated with the image of the jth basis vector $\mathbf{x}_j$.

If we change to a different basis, the matrix of A will change. To see how, suppose $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ and $\{\widetilde{\mathbf{x}}_1, \dots, \widetilde{\mathbf{x}}_n\}$ are bases for V, and let $C = (C_{ij})$ be the matrix of coefficients of $\widetilde{\mathbf{x}}_j$ expressed with respect to the basis $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$: thus for each $j = 1, \dots, n$,

$$\widetilde{\mathbf{x}}_j = \sum_{i=1}^{n} C_{ij}\mathbf{x}_i. \tag{3.1}$$

The matrix C is called the ***transition matrix*** from $\{\mathbf{x}_i\}$ to $\{\widetilde{\mathbf{x}}_j\}$. Its columns are the n-tuples representing $\widetilde{\mathbf{x}}_1, \dots, \widetilde{\mathbf{x}}_n$ in terms of the basis $\{\mathbf{x}_i\}$, which are linearly independent, so C is invertible.

Proposition 3.1 (Change of Basis Formula). *Suppose V is a finite-dimensional vector space and $A\colon V \to V$ is an endomorphism. Given two bases $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ and $\{\widetilde{\mathbf{x}}_1, \dots, \widetilde{\mathbf{x}}_n\}$ for V, the matrices $A_{\mathbf{x}}$ and $A_{\widetilde{\mathbf{x}}}$ representing A with respect to the two bases are related by*

$$A_{\mathbf{x}} = C A_{\widetilde{\mathbf{x}}} C^{-1}, \tag{3.2}$$

where C is the transition matrix defined by (3.1)*.*

Proof. Let (A_{ij}) denote the matrix entries of $A_{\mathbf{x}}$ and $(\widetilde{A}_{ij})$ those of $A_{\widetilde{\mathbf{x}}}$. We prove the proposition by calculating the vector $A\widetilde{\mathbf{x}}_j$ in two ways. First, we substitute (3.1) and then expand $A\mathbf{x}_i$ in terms of (A_{ij}):

$$A\widetilde{\mathbf{x}}_j = \sum_{i=1}^{n} C_{ij} A\mathbf{x}_i = \sum_{i,k=1}^{n} C_{ij} A_{ki} \mathbf{x}_k.$$

Second, we expand $A\widetilde{\mathbf{x}}_j$ in terms of $(\widetilde{A}_{ij})$ and then substitute (3.1):

$$A\widetilde{\mathbf{x}}_j = \sum_{i=1}^{n} \widetilde{A}_{ij} \widetilde{\mathbf{x}}_i = \sum_{i,k=1}^{n} \widetilde{A}_{ij} C_{ki} \mathbf{x}_k.$$

Because the vectors $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ are independent, the fact that these two expressions are equal implies that the respective coefficients of $\mathbf{x}_k$ are equal:

$$\sum_{i=1}^{n} A_{ki} C_{ij} = \sum_{i=1}^{n} C_{ki} \widetilde{A}_{ij}.$$

This is equivalent to the matrix equation $A_{\mathbf{x}} C = C A_{\widetilde{\mathbf{x}}}$, which in turn is equivalent to (3.2). $\square$

Here is the most important application of the change of basis formula. You have already seen the ***determinant*** of an $n \times n$ matrix (see Handout 1). The ***trace*** of an $n \times n$ matrix M is the number $\operatorname{tr} M = \sum_i M_{ii}$ (the sum of the entries on the main diagonal). The next theorem describes some of the most important properties of the determinant and trace functions.

Theorem 3.2. *For any $n \times n$ matrices M and N,*

$$\det(MN) = (\det M)(\det N) = \det(NM); \tag{3.3}$$

$$\operatorname{tr}(MN) = \operatorname{tr}(NM). \tag{3.4}$$

Proof. For a proof of (3.3), see any good linear algebra book. For (3.4), we compute as follows:

$$\operatorname{tr}(MN) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} M_{ij} N_{ji} \right) = \sum_{i,j=1}^{n} M_{ij} N_{ji};$$

and $\operatorname{tr}(NM)$ yields the same expression with the roles of i and j reversed. $\square$

Corollary 3.3. *Suppose V is a finite-dimensional vector space and $A\colon V \to V$ is an endomorphism. If $\{\mathbf{x}_i\}$ and $\{\widetilde{\mathbf{x}}_j\}$ are any two bases for V and $A_{\mathbf{x}}$ and $A_{\widetilde{\mathbf{x}}}$ are the matrix representations of A with respect to the two bases, then $\det A_{\mathbf{x}} = \det A_{\widetilde{\mathbf{x}}}$ and $\operatorname{tr} A_{\mathbf{x}} = \operatorname{tr} A_{\widetilde{\mathbf{x}}}$.*

Proof. Let C be the transition matrix from $\{\mathbf{x}_i\}$ to $\{\widetilde{\mathbf{x}}_j\}$. Using the results of Proposition 3.1 and Theorem 3.2, we compute

$$\begin{aligned}
\det A_{\mathbf{x}} &= \det\left(C(A_{\widetilde{\mathbf{x}}}C^{-1})\right)\\
&= \det\left((A_{\widetilde{\mathbf{x}}}C^{-1})C\right)\\
&= \det\left(A_{\widetilde{\mathbf{x}}}(C^{-1}C)\right)\\
&= \det A_{\widetilde{\mathbf{x}}}.
\end{aligned}$$

The computation for the trace is identical. □

Because of this corollary, we can make the following definition: if $A\colon V \to V$ is any endomorphism, we define the ***determinant of A*** to be the determinant of any matrix representation of A, and the ***trace of A*** to be the trace of any matrix representation. The corollary shows that these numbers are well defined, independently of the choice of basis.

Bilinear Forms

A ***bilinear form*** on V is a function $B\colon V \times V \to \mathbb{R}$ that is linear in each variable separately; this means that for all $v, w, x \in V$ and all $a, b \in \mathbb{R}$ it satisfies

$$\begin{aligned}
B(av + bw, x) &= aB(v, x) + bB(w, x),\\
B(x, av + bw) &= aB(x, v) + bB(x, w).
\end{aligned}$$

A bilinear form B is said to be ***symmetric*** if $B(v, w) = B(w, v)$ for all $v, w \in V$, and it is said to be ***positive definite*** if $B(v, v) \geq 0$ for all $v \in V$, with equality if and only if $v = 0$.

It is important to see what bilinear forms look like in terms of a basis. Let B be a bilinear form on V and suppose $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ is a basis for V. We define an $n \times n$ matrix $B_{\mathbf{x}} = (B_{ij})$, called the ***matrix of B with respect to this basis***, by

$$B_{ij} = B(\mathbf{x}_i, \mathbf{x}_j)$$

Then because B is bilinear, its action on any pair of vectors $v = \sum_i v_i\mathbf{x}_i$ and $w = \sum_j w_j\mathbf{x}_j$ can be computed as follows:

$$B(v, w) = B\left(\sum_{i=1}^n v_i\mathbf{x}_i,\ \sum_{j=1}^n w_j\mathbf{x}_j\right) = \sum_{i,j=1}^n v_i w_j B(\mathbf{x}_i, \mathbf{x}_j) = \sum_{i,j} B_{ij} v_i w_j. \tag{3.5}$$

This can be summarized as the value obtained by multiplying the matrix $B_{\mathbf{x}}$ on the right by w and on the left by the transpose of v:

$$B(v, w) = \begin{pmatrix} v_1 & \dots & v_n \end{pmatrix} \begin{pmatrix} B_{11} & \dots & B_{1n}\\ \vdots & \ddots & \vdots\\ B_{n1} & \dots & B_{nn} \end{pmatrix} \begin{pmatrix} w_1\\ \vdots\\ w_n \end{pmatrix}. \tag{3.6}$$

In matrix notation, we can write this as $B(v, w) = v_{\mathbf{x}}^T B_{\mathbf{x}} w_{\mathbf{x}}$, where $v_{\mathbf{x}}$ and $w_{\mathbf{x}}$ are the column matrices representing v and w in this basis, and the superscript T designates the ***transpose*** of a matrix: if $M = (M_{ij})$ is any $k \times l$ matrix, its transpose M^T is the $l \times k$ matrix whose (i, j)-entry is $(M^T)_{ij} = M_{ji}$. In particular, (3.5) implies that if two bilinear forms agree on all pairs of vectors in some basis, then they are identical.

A matrix is said to be ***symmetric*** if it is equal to its transpose. Note that if a bilinear form B is symmetric, then its matrix with respect to any basis is a symmetric matrix, because

$$B_{ij} = B(\mathbf{x}_i, \mathbf{x}_j) = B(\mathbf{x}_j, \mathbf{x}_i) = B_{ji}.$$

Conversely, if B is represented by a symmetric matrix with respect to some basis, then it follows easily from (3.5) that it is a symmetric bilinear form.

The most important type of bilinear form is an ***inner product***, which is a bilinear form that is symmetric and positive definite. Given an inner product on V, we usually denote the value of the inner product at a pair of vectors v and w by the notation $\langle v, w\rangle$. A ***(finite-dimensional) inner product space*** is a finite-dimensional vector space endowed with a specific choice of inner product. The most familiar and important example is $\mathbb{R}^n$ with its Euclidean dot product,

$$\langle v, w\rangle = v \cdot w = v_1 w_1 + \cdots + v_n w_n.$$

The following exercise shows a common way to construct other examples.

► **Exercise 3.4.** Suppose $(V, \langle\,\cdot\,,\,\cdot\,\rangle)$ is a finite-dimensional inner product space and $W \subseteq V$ is a linear subspace of V. Prove that the restriction of $\langle\,\cdot\,,\,\cdot\,\rangle$ to $W \times W$ is an inner product on W.

Henceforth, we assume that V is an n-dimensional inner product space, endowed with a specific inner product $\langle\,\cdot\,,\,\cdot\,\rangle$. (In our applications, V will be a tangent plane to a surface, and $\langle\,\cdot\,,\,\cdot\,\rangle$ will be the restriction of the Euclidean dot product.)

In an inner product space, we can define many geometric quantities analogous to ones that we are familiar with in $\mathbb{R}^n$. For example, the ***norm*** of a vector $v \in V$ is the nonnegative real number $\|v\| = \langle v, v\rangle^{1/2}$, and the ***angle*** between two nonzero vectors v, w is $\theta = \arccos\big(\langle v, w\rangle/(\|v\|\,\|w\|)\big)$. A ***unit vector*** is a vector v with $\|v\| = 1$, and two vectors v, w are ***orthogonal*** if $\langle v, w\rangle = 0$. A set of vectors $\{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_k\}$ is said to be ***orthonormal*** if each $\boldsymbol{\varepsilon}_i$ is a unit vector and distinct vectors are orthogonal; or, more succinctly, if

$$\langle \boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_j\rangle = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

(The symbol δ_{ij} is called the ***Kronecker delta***.) An ***orthonormal basis*** is a basis consisting of orthonormal vectors.

Lemma 3.5. *If $\{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_k\}$ is a set of orthonormal vectors, then it is a linearly independent set.*

Proof. Given an orthonormal set $\{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_k\}$, suppose $0 = a_1\boldsymbol{\varepsilon}_1 + \cdots + a_k\boldsymbol{\varepsilon}_k$. Taking the inner product of both sides with $\boldsymbol{\varepsilon}_i$, we find that $0 = a_i\langle\boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_i\rangle = a_i$. Thus all of the coefficients $a_1, \dots, a_k$ are zero. □

The next proposition shows that every inner product space admits many orthonormal bases.

Proposition 3.6 (Gram–Schmidt Algorithm). *Let V be an inner product space and let $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ be any basis for V. Then there exists an orthonormal basis $\{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_n\}$ such that $\operatorname{span}(\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_k) = \operatorname{span}(\mathbf{x}_1, \dots, \mathbf{x}_k)$ for each $k = 1, \dots, n$.*

Proof. We will prove by induction on k that for each $k = 1, \dots, n$ there exist orthonormal vectors $\{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_k\}$ whose span is the same as that of $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$. When $k = n$, this proves the proposition, because orthonormal vectors are independent, and a linearly independent set of n vectors in an n-dimensional vector space is automatically a basis.

Begin by setting $\boldsymbol{\varepsilon}_1 = \mathbf{x}_1/\|\mathbf{x}_1\|$, which is a unit vector whose span is the same as that of $\mathbf{x}_1$. Now let $k \geq 1$ and assume by induction that we have produced orthonormal vectors $\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_k$ satisfying the span condition. Define

$$\mathbf{y}_{k+1} = \mathbf{x}_{k+1} - \sum_{i=1}^{k} \langle \mathbf{x}_{k+1}, \boldsymbol{\varepsilon}_i \rangle \boldsymbol{\varepsilon}_i,$$
$$\boldsymbol{\varepsilon}_{k+1} = \frac{\mathbf{y}_{k+1}}{\|\mathbf{y}_{k+1}\|}.$$

Because $\mathbf{x}_{k+1} \notin \operatorname{span}(\mathbf{x}_1, \dots, \mathbf{x}_k)$, it follows that $\mathbf{y}_{k+1} \neq 0$, and thus $\boldsymbol{\varepsilon}_{k+1}$ is well defined. Clearly $\boldsymbol{\varepsilon}_{k+1}$ is a unit vector. A straightforward computation shows that $\mathbf{y}_{k+1}$ is orthogonal to each of the vectors $\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_k$, and therefore so is $\boldsymbol{\varepsilon}_{k+1}$. Since $\boldsymbol{\varepsilon}_{k+1}$ is a linear combination of the vectors $\{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_k, \mathbf{x}_{k+1}\}$, it lies in their span, which by the induction hypothesis is equal to $\operatorname{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{x}_{k+1}\}$. Since the vectors $\{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_{k+1}\}$ are orthonormal, they are independent, and thus their span is a $(k+1)$-dimensional subspace contained in the span of $\{\mathbf{x}_1, \dots, \mathbf{x}_{k+1}\}$. These two subspaces have the same dimension, so they are equal. $\square$

Like any symmetric bilinear form, an inner product can be represented in terms of a basis by a symmetric matrix. It is traditional to write the matrix of an inner product as $g_{\mathbf{x}} = (g_{ij})$, where $g_{ij} = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$. In an orthonormal basis, the inner product is represented by the identity matrix, but in terms of a non-orthonormal base it will be represented by a different matrix.

Bilinear Forms on an Inner Product Space

We continue to assume that V is an n-dimensional inner product space. By means of the inner product, we can construct many bilinear forms on V as follows. Given any endomorphism $A\colon V \to V$, we define a bilinear form B_A on V by the following formula:

$$B_A(v, w) = \langle v, Aw \rangle. \tag{3.7}$$

► **Exercise 3.7.** Prove that B_A is in fact a bilinear form.

In fact, this example is not special, because, as the following theorem shows, every bilinear form can be constructed in this way.

Theorem 3.8. *Let V be a finite-dimensional inner product space, and let B be a bilinear form on V. Then there exists a unique endomorphism $A\colon V \to V$ such that $B = B_A$. In terms of any orthonormal basis for V, A and B are represented by the same matrix.*

Proof. Let $\{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_n\}$ be any orthonormal basis for V, and write $B_{ij} = B(\boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_j)$. Let $A\colon V \to V$ be the endomorphism determined by the same matrix with respect to this basis, so that

$$A\boldsymbol{\varepsilon}_j = \sum_{k=1}^{n} B_{kj} \boldsymbol{\varepsilon}_k.$$

For each i, j, we compute

$$B_A(\boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_j) = \langle \boldsymbol{\varepsilon}_i, A\boldsymbol{\varepsilon}_j \rangle = \sum_{k=1}^{n} B_{kj} \langle \boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_k \rangle = B_{ij},$$

where the last equation follows because the only term in the sum for which $\langle \boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_k \rangle \neq 0$ is the one with $k = i$. Thus B_A and B give the same results when applied to pairs of basis vectors, so they are equal.

To prove uniqueness, suppose A_1 and A_2 are endomorphisms such that $\langle v, A_1 w \rangle = B(v,w) = \langle v, A_2 w \rangle$ for all $v, w \in V$. Define $D\colon V \to V$ by $Dw = A_1 w - A_2 w$. The hypothesis implies that $\langle v, Dw \rangle = 0$ for all $v, w \in V$. In particular, taking $v = Dw$, this implies $0 = \langle Dw, Dw \rangle = \|Dw\|^2$ for all $w \in V$. Thus D is the zero endomorphism, which implies $A_1 = A_2$. $\square$

Given a bilinear form B, the unique endomorphism A such that $B = B_A$ is called the ***endomorphism associated with B***. Similarly, given an endomorphism A, we say that the bilinear form B_A defined by (3.7) is ***associated with A***. Note that the endomorphism associated with a bilinear form is canonically determined, independent of any choice of basis, even though we used a basis to prove its existence and uniqueness.

It is important to be aware that a bilinear form and its associated endomorphism are represented by the same matrix *only* when working with an orthonormal basis. The next proposition shows how the matrices are related in an arbitrary basis.

Proposition 3.9. *Suppose B is a bilinear form on V and A is its associated endomorphism. In terms of any basis $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, the matrices of B and A are related by $B_{\mathbf{x}} = g_{\mathbf{x}} A_{\mathbf{x}}$, where $g_{\mathbf{x}}$ is the matrix representing the inner product.*

Proof. If v and w are arbitrary vectors in V, then

$$v_{\mathbf{x}}^T B_{\mathbf{x}} w_{\mathbf{x}} = B(v,w) = \langle v, Aw \rangle = v_{\mathbf{x}}^T g_{\mathbf{x}} A_{\mathbf{x}} w_{\mathbf{x}}.$$

This shows that the matrices $B_{\mathbf{x}}$ and $g_{\mathbf{x}} A_{\mathbf{x}}$ represent the same bilinear form. Since the (i,j)-entry of such a matrix is the result of evaluating the bilinear form on $(\mathbf{x}_i, \mathbf{x}_j)$, the two matrices are equal. $\square$

We will be concerned primarily with symmetric bilinear forms, so it is important to identify a property of endomorphisms that corresponds to symmetry. If $A\colon V \to V$ is an endomorphism, we say that A is ***symmetric*** or ***selfadjoint*** if the following identity holds for all $v, w \in V$:

$$\langle v, Aw \rangle = \langle Av, w \rangle.$$

It is obvious from the definition that a bilinear form is symmetric if and only if its associated endomorphism is symmetric. It follows from Theorem 3.8 that an endomorphism is symmetric if and only if its matrix with respect to any *orthonormal* basis is symmetric. (Be warned, however, that the matrix of a symmetric endomorphism with respect to a non-orthonormal basis might *not* be symmetric.)

Note that although symmetry of a bilinear form makes sense without reference to an inner product, it only makes sense to say that an endomorphism is symmetric if a specific inner product has been chosen.

If $A\colon V \to V$ is an endomorphism, a real number λ is called an ***eigenvalue of A*** if there exists a nonzero vector $v \in V$ such that $Av = \lambda v$. Any such vector is called an ***eigenvector of A*** corresponding to λ. (In many applications of linear algebra, it is necessary also to consider complex eigenvalues, which require a slightly more elaborate definition; but real eigenvalues will suffice for our purposes.)

The most important fact about symmetric endomorphisms is the following theorem.

Theorem 3.10 (Finite-Dimensional Spectral Theorem). *Suppose V is an inner product space and $A\colon V \to V$ is a symmetric endomorphism. Then V has an orthonormal basis consisting of eigenvectors of A.*

Proof. We will only need the theorem when V is 2-dimensional, so we give the proof for that case only. Choose any orthonormal basis $\{\mathbf{x}_1, \mathbf{x}_2\}$ for V. First we dispose of a simple special case: if A is equal to a scalar multiple of the identity map, meaning that there is a real number λ such that $Ax = \lambda x$ for all $x \in V$, then the chosen basis $\{\mathbf{x}_1, \mathbf{x}_2\}$ is already an orthonormal basis of eigenvectors and we are done. So assume henceforth that A is not a scalar multiple of the identity.

With respect to the chosen orthonormal basis, A is represented by a symmetric matrix:

$$A_{\mathbf{x}} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

A real number λ is eigenvalue of A if and only if there is a nonzero vector v such that $Av - \lambda v = 0$. This is the case if and only if the matrix $A_{\mathbf{x}} - \lambda I$ is singular (where I is the 2×2 identity matrix), which is equivalent to $\det(A_{\mathbf{x}} - \lambda I) = 0$. Thus the eigenvalues of A are the solutions (if any) to the following quadratic equation:

$$(a - \lambda)(c - \lambda) - b^2 = 0,$$

or equivalently

$$\lambda^2 - (a + c)\lambda + (ac - b^2) = 0.$$

This has solutions

$$\lambda = \frac{a + c \pm \sqrt{(a + c)^2 - 4(ac - b^2)}}{2} = \frac{a + c \pm \sqrt{(a - c)^2 + 4b^2}}{2}.$$

Since we are assuming that A is not a multiple of the identity, it must be the case that either $a \neq c$ or $b \neq 0$; in either case the expression under the square root sign is strictly positive, so the quadratic equation has two distinct real roots. Call them λ_1 and λ_2.

For each $j = 1, 2$, the fact that $A - \lambda_j \operatorname{id}$ is singular means it has nontrivial kernel, so there exist nonzero vectors $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ such that

$$\begin{aligned} A\boldsymbol{\varepsilon}_1 - \lambda_1 \boldsymbol{\varepsilon}_1 &= 0, \\ A\boldsymbol{\varepsilon}_2 - \lambda_2 \boldsymbol{\varepsilon}_2 &= 0. \end{aligned}$$

After dividing each vector $\boldsymbol{\varepsilon}_j$ by its norm (which does not affect the two equations above), we may assume that $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ are unit vectors.

Finally, we will show that $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ are orthogonal. Using the fact that A is symmetric, we compute

$$\lambda_1 \langle \boldsymbol{\varepsilon}_1,\ \boldsymbol{\varepsilon}_2 \rangle = \langle \lambda_1 \boldsymbol{\varepsilon}_1,\ \boldsymbol{\varepsilon}_2 \rangle = \langle A\boldsymbol{\varepsilon}_1,\ \boldsymbol{\varepsilon}_2 \rangle = \langle \boldsymbol{\varepsilon}_1,\ A\boldsymbol{\varepsilon}_2 \rangle = \langle \boldsymbol{\varepsilon}_1,\ \lambda_2 \boldsymbol{\varepsilon}_2 \rangle = \lambda_2 \langle \boldsymbol{\varepsilon}_1,\ \boldsymbol{\varepsilon}_2 \rangle.$$

Thus $(\lambda_1 - \lambda_2)\langle \boldsymbol{\varepsilon}_1,\ \boldsymbol{\varepsilon}_2 \rangle = 0$, and since $\lambda_1 \neq \lambda_2$, this implies $\langle \boldsymbol{\varepsilon}_1,\ \boldsymbol{\varepsilon}_2 \rangle = 0$. □

Quadratic Forms

In many applications, the most important uses of bilinear forms involve their values when both arguments are the same. For that reason, we make the following definition. If V is a finite-dimensional vector space, a function $Q\colon V \to \mathbb{R}$ is called a ***quadratic form on V*** if there is some bilinear form B on V such that $Q(v) = B(v, v)$ for all $v \in V$. Any such bilinear form is said to be ***associated with Q***. The next proposition shows that if we require B to be symmetric, it is uniquely determined by Q.

Proposition 3.11. *If Q is any quadratic form on a finite-dimensional vector space V, there is a unique symmetric bilinear form associated with Q.*

Proof. Given a quadratic form Q on V, by definition there is some bilinear form B_0 such that $Q(v) = B_0(v,v)$ for all v. Define $B\colon V \times V \to \mathbb{R}$ by

$$B(v,w) = \tfrac{1}{2}\big(B_0(v,w) + B_0(w,v)\big).$$

Then an easy verification shows that B is bilinear, and it is obviously symmetric. For any $v \in V$, we have

$$B(v,v) = \tfrac{1}{2}\big(B_0(v,v) + B_0(v,v)\big) = B_0(v,v) = Q(v),$$

so this proves the existence of such a B.

To prove uniqueness, we will derive a formula (called a ***polarization identity***) which shows that B is completely determined by Q. If B is any symmetric bilinear form associated with Q, we have

$$\begin{aligned}
\tfrac{1}{4}\big(Q(v+w) - Q(v-w)\big) &= \tfrac{1}{4}\big(B(v+w, v+w) - B(v-w, v-w)\big)\\
&= \tfrac{1}{4}\big(B(v,v) + B(w,v) + B(v,w) + B(w,w)\big)\\
&\quad - \tfrac{1}{4}\big(B(v,v) - B(w,v) - B(v,w) + B(w,w)\big)\\
&= B(v,w).
\end{aligned} \tag{3.8}$$

Any other symmetric bilinear form associated with Q would have to satisfy an analogous equation, and thus would be equal to B. $\square$