Notes and Exercises on Bilinear Forms

Instructor: K N Raghavan

AFS-I at MEPCO 2025, second week, 8th to 12th December 2025

Bilinear and Sesqui-linear Forms: Definitions and Basic Properties

Let k be any field, V be a finite dimensional vector space over k, and put $n := \dim_k V$. A *bilinear form* on V is a map $V \times V \to k$, denoted $(v, w) \mapsto \langle v, w\rangle$, that is linear in both arguments, or, more precisely, that satisfies these two conditions, for all v, v', w and w' in V and λ, μ in k:

$$\langle \lambda v + \mu v', w\rangle = \lambda \langle v, w\rangle + \mu \langle v', w\rangle \quad \text{and} \quad \langle v, \lambda w + \mu w'\rangle = \lambda \langle v, w\rangle + \mu \langle v, w'\rangle$$

The standard dot product on $\mathbb{R}^n$ is a bilinear form: identifying $\mathbb{R}^n$ with $n \times 1$ column matrices over $\mathbb{R}$, the dot product of two vectors x, y in $\mathbb{R}^n$ is defined by $x \cdot y := x^t y$ (where $x^t y$ denotes the matrix product of the transpose of x with y).

The standard dot product on $\mathbb{C}^n$ is however <u>not</u> a bilinear form: identifying $\mathbb{C}^n$ with $n \times 1$ column matrices over $\mathbb{C}$, the dot product of two vectors w, z in $\mathbb{C}^n$ is defined by $w \cdot z := w^\star z$ (where $w^\star z$ denotes the matrix product of the conjugate-transpose of w with z). Note that $w^\star z$ is not linear but only “conjugate-linear” in its first argument w. The following definition is tailored to capture the complex dot product.

Put $k = \mathbb{C}$. A *sesqui-linear form* on V is a map $V \times V \to \mathbb{C}$ that satisfies the following two conditions for all v, v', w, and w' in V, and λ, μ in $\mathbb{C}$:

$$\langle \lambda v + \mu v', w\rangle = \overline{\lambda} \langle v, w\rangle + \overline{\mu} \langle v', w\rangle \quad \text{and} \quad \langle v, \lambda w + \mu w'\rangle = \lambda \langle v, w\rangle + \mu \langle v, w'\rangle$$

The real dot product is “symmetic”, that is, $x \cdot y = y \cdot x$; the complex dot product, on the other hand, is “conjugate-symmetric”: $w \cdot z = \overline{z \cdot w}$. Both these products are “positive”: $x \cdot x \geq 0$ for all x, and $z \cdot z \geq 0$ for all z (note that the conjugate-symmetry property implies that $z \cdot z$ is a real number). And they are both “definite”: $x \cdot x = 0$ only if $x = 0$; and $z \cdot z = 0$ only if $z = 0$.

Inspired by these qualities of the real and complex dot product, we make the following definitions:

- a bilinear form (k is now any field) is *symmetric* if $\langle v, w\rangle = \langle w, v\rangle$ for all v and w in V.
- (with $k = \mathbb{C}$) a sesqui-linear form is *Hermitian* if $\langle v, w\rangle = \overline{\langle w, v\rangle}$; *skew-Hermitian* if $\langle v, w\rangle = -\overline{\langle w, v\rangle}$.
- Let $\langle\ ,\ \rangle$ be a symmetric bilinear form with $k = \mathbb{R}$ or a Hermitian (sesqui-linear) form with $k = \mathbb{C}$;
  - it is said to be *positive semi-definite* or sometimes just *positive* if $\langle v, v\rangle \geq 0$;
  - it is said to be *negative semi-definite* or sometimes just *negative* if $\langle v, v\rangle \leq 0$;
  - it is said to be *indefinite* if it is neither positive nor negative.
- a positive or negative (bilinear or sequi-linear) form is *definite* if $\langle v, v\rangle = 0$ implies $v = 0$.

The real dot product is symmetric and positive definite; the complex dot product is Hermitian and positive definite. When $k = \mathbb{C}$, we shall consider both bilinear and sesqui-linear forms.

Exercises

1. Observe that $\langle 0, 0\rangle = 0$ for any (bilinear or sesqui-linear) form $\langle\ ,\ \rangle$.
2. Each of the following is a symmetric form on $\mathbb{R}^2$: x_1, x_2 denote the co-ordinates of the vector x (and similarly for y). Determine if it is positive, negative, or indefinite; and in case it is positive or negative, whether or not it is definite. (a) $\langle x, y\rangle = -x_1y_1 - x_2y_2$; (b) $\langle x, y\rangle = x_1y_1 - x_2y_2$; (c) $\langle x, y\rangle = x_1y_1 - x_2y_2$; (d) $\langle x, y\rangle = x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$; (e) $\langle x, y\rangle = x_1y_1 - \frac{1}{2}x_1y_2 - \frac{1}{2}x_2y_1 + x_2y_2$; (f) $\langle x, y\rangle = x_1y_1 + 2x_1y_2 + 2x_2y_1 + x_2y_2$.
3. (Representation by bilinear polynomials) Consider polynomials of the form $\sum_{i,j} a_{ij}x_iy_j$ in two sets of n variables $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ with coefficients in a field k. Such a polynomial defines a bilinear form on k^n by $\langle x, y\rangle := \sum_{i,j} a_{ij}x^iy^j$ (where x_i and y_j are respectively the co-ordinates of x and y). We have $\langle e_i, e_j\rangle = a_{ij}$ (where $e_1, \ldots, e_n$ are the standard basis vectors of k^n). Every bilinear form is defined by such a polynomial: the form $\langle\ ,\ \rangle$ comes from $\sum_{i,j} \langle e_i, e_j\rangle\, x_iy_j$. Evidently, a form is symmetric if and only if the corresponding polynomial is unchanged if we exchange x_i with y_i (for all $1 \leq i \leq n$).

Matrix Representation of Bilinear and Sesqui-linear Forms

Let V be a finite dimensional vector space over an arbitrary field k. Put $n := \dim_k V$. Let us fix a basis $\mathcal{B} := \{b_1, \ldots, b_n\}$ of V. Given a bilinear form $\langle\ ,\ \rangle$, we can "represent" it, using our fixed basis, by means of an $n \times n$ matrix, which we denote $A_{\mathcal{B}}$: the $(i,j)^{\text{th}}$ entry of $A_{\mathcal{B}}$ is $\langle b_i, b_j\rangle$. Conversely, given any $n \times n$ matrix A, we can define a form $\langle\ ,\ \rangle_A$ on V, whose matrix representation with respect to $\mathcal{B}$ is A: $\langle v, w\rangle := \sum_{i,j} a_{ij} x_i y_j$, where the x_i and y_j are the co-ordinates of v and w with respect to $\mathcal{B}$: $v = \sum x_i b_i$ and $w = \sum_j y_j b_j$. If, as is usual, we put the x_i (respectively y_j) together to form a $n \times 1$ matrix, we can write $\langle v, w\rangle = x^t A y$. This sets up a one-to-one correspondence, depending on the chosen basis $\mathcal{B}$, between bilinear forms on V and $n \times n$ matrices A. This correspondence restricts to give a one-to-one correspondence between symmetric bilinear forms and symmetric matrices.

1. (How does the matrix of a form change with basis?) Let $\langle\ ,\ \rangle$ be a fixed bilinear form on an n-dimensional vector space V (over an arbitrary field k). Let $A_{\mathcal{B}}$ and $A_{\mathcal{B}'}$ be the matrices of the form with respect to two bases $\mathcal{B}$ and $\mathcal{B}'$ of V. Work out the relationship between the two matrices.

   Solution: Let P be the invertible $n \times n$ matrix such that $\mathcal{B}' = \mathcal{B}P$. Let x, x' be the co-ordinates of a vector v with respect to $\mathcal{B}$ and $\mathcal{B}'$ respectively. Similarly, let y, y' be the co-ordinate of a vector w with respect to $\mathcal{B}$ and $\mathcal{B}'$ respectively. We have the following relations: $Px' = x$, $Py' = y$. Using these to substitute for x and y in $\langle v, w\rangle = x^t A_{\mathcal{B}} y$, we get $\langle v, w\rangle = x'^t P^t A_{\mathcal{B}} P y'$. But, on the other hand, $\langle v, w\rangle = x'^t A_{\mathcal{B}'} y'$, and so $x'^t P^t A_{\mathcal{B}} P y' = x'^t A_{\mathcal{B}'} y'$. Since this equality holds for all x' and y' in k^n, we conclude that $A_{\mathcal{B}'} = P^t A_{\mathcal{B}} P$. □

2. The set of all bilinear forms on an n-dimensional vector space V (over an arbitrary field k) is naturally a k-vector space: the addition and scalar multiplication are defined in the obvious way. The one-to-one correspondence between forms and matrices (described above) with respect to a fixed basis is a vector space isomorphism. In particular, the dimension of the space of all bilinear forms on V is n^2. What is the dimension of the subspace of symmetric forms?

   Solution: Symmetric forms correspond to symmetric matrices (under the isomorphism), and so form a vector space of dimension $n(n+1)/2$. □

3. Let V be an n-dimensional complex vector space (for some non-negative integer n), and fix a basis $\mathcal{B}$ for it. Proceeding as in the case of bilinear forms, we get a bijective correspondence between sesqui-linear forms on V and $n \times n$ complex matrices (defined with respect to the fixed basis $\mathcal{B}$), except now the formula for the form in terms of the matrix becomes: $\langle v, w\rangle = \sum_{i,j} a_{ij} \overline{x_i} y_j = x^\star A y$. (Here and elsewhere $\star$ denotes conjugate-transpose.)

4. Observe that the set of sesqui-linear forms, with addition and scalar multiplication defined in the obvious way, forms a real vector space (although not a complex vector space). The subset of Hermitian forms (respectively skew-Hermitian forms) is a subspace. The correspondence above is an $\mathbb{R}$-vector space isomorphism between the space of sesqui-linear forms and $n \times n$ complex matrices. It restricts to an isomorphism between the subspace of Hermitian forms (respectively skew-Hermitian forms) and the subspace of Hermitian matrices (respectively skew-Hermitian matrices). (Recall that a complex $n \times n$ matrix A is called *Hermitian* if $A = A^\star$ and *skew-Hermitian* if $A = -A^\star$.) What are the dimensions of the spaces of Hermitian forms and skew-Hermitian forms?

   Solution: The dimension is n^2 of all these $\mathbb{R}$-vector spaces: Hermitian forms on n-dimensional complex vector space V; $n \times n$ Hermitian matrices; skew-Hermitian forms on V; $n \times n$ skew-Hermitian matrices. The first two are isomorphic to each other and the last two are isomorphic to each other as we have just seen. Multiplication by the complex number i sets up an isomorphism between the spaces of Hermitian and skew-Hermitian matrices (respectively forms). □

5. Show that every sesqui-linear form on a finite dimensional complex vector space is uniquely a sum of a Hermitian form and a skew-Hermitian form.

   Solution: Choosing a basis for the vector space, we get an isomorphism of $\mathbb{R}$-vector spaces between sesqui-linear forms and complex $n \times n$ matrices (where n is the dimension of the vector space). Every $n \times n$ matrix A is uniquely the sum of a Hermitian and skew-Hermitian matrix: $A = (A + A^\star)/2 + (A - A^\star)/2$. Pulling this back to forms via the isomorphism, we get the desired result. □

Orthogonality, Null Space, and Non-degeneracy

Let V be a finite dimensional vector space over a field k. Put $n := \dim_k V$. Let $\langle\ ,\ \rangle$ be a symmetric bilinear form on V or a Hermitian form (in the latter case, $k = \mathbb{C}$).

1. Given vectors v, w in V, we call them *orthogonal* or *perpendicular* and write $v \perp w$ if $\langle v, w\rangle = 0$. **Caveat**: One should be wary of transferring intuition about orthogonality in $\mathbb{R}^2$ and $\mathbb{R}^3$ to this more general context. For example, we could consider an $n \times n$ symmetric matrix A with zeros on the diagonal—there are plenty of such matrices—and use A, by fixing a basis of V, to define a bilinear form on V. For any such form, $v \perp v$ for all vectors v!

2. Given a vector v_0 in V, the assignment $v \mapsto < v_0, v >$ defines a linear functional on V, and its kernel has dimension at least $n - 1$. If this functional is identically zero, then v_0 is said to be in the *null space* of the form.

3. For W a subspace of V, define $W^\perp := \{v \in V \mid \langle v, w\rangle = 0\, \forall w \in W\}$. Evidently, $W^\perp$ is a subspace of V, and $(W^\perp)^\perp \supseteq W$. Show that $\dim W^\perp \geq \dim V - \dim W$.

   Solution: Put $m := \dim W$ and let $w_1, \ldots, w_m$ be a basis of W. Define a linear map $T : V \to k^m$ by $v \mapsto (\langle v, w_1\rangle, \langle v, w_2\rangle, \ldots, \langle v, w_m\rangle)$. The kernel of the map T is $W^\perp$. By the rank-nullity theorem, we get $\dim W^\perp = \dim V - \operatorname{rank} T \geq \dim V - m = \dim V - \dim W$. □

4. The *null space* of the form, denoted $V^\perp$, is defined by $V^\perp := \{v \in V \mid \langle v, w\rangle = 0\ \forall\ w \in V\}$. Let A be a matrix of the form, with respect to some basis. Show that an element v in V belongs to the null space if and only if $Ax = 0$ where x is the $n \times 1$ matrix of co-ordinates of v with respect to the same basis.

   Solution: We have $\langle v, w\rangle = x^t Ay$, where x and y are the co-ordinate matrices of v and w (with respect to the fixed basis). For v to be in the null space, it means that $x^t Ay = 0$ for all possible $y \in \mathbb{R}^n$, but this happens if and only if $x^t A = 0$. By taking transpose and using the fact that A is symmetric (since the form is symmetric), we get $Ax = 0$. □

5. Our fixed symmetric / Hermitian form $\langle\ ,\ \rangle$ is said to be *non-degenerate* if its null space is zero. Show that this happens if and only if the matrix A of the form with respect to some basis is invertible.

   Solution: Since matrices with respect to two different basis are related as A is to $P^t AP$ for some invertible matrix P, it is clear that the invertibility of the matrix A is independent of the choice of basis. To prove the assertion, we deduce, using the previous item, that v belongs to the null space if and only if $Ax = 0$ (where x denotes the matrix of co-ordinates of the vector v with respect to the fixed basis). But $Ax = 0$ has a non-trivial solution if and only if A is not invertible. □

Restricting Forms to Subspaces

Let V be a finite dimensional vector space over a field k. Put $n := \dim_k V$. Let $\langle\ ,\ \rangle$ be a symmetric bilinear form on V or a Hermitian form (in the latter case, $k = \mathbb{C}$). The form restricts to any subspace W of V to give a similar form on W. The restricted form inherits any of the following properties of the original form on V (in case we can speak about them): positive definiteness, positive semi-definiteness, negative definiteness, negative semi-definiteness. However an indefinite form could restrict to a definite one; a non-degenerate form could restrict to a degenerate one; and a degenerate one could restrict to a non-degenerate one.

6. (Lemma 8.4.2 of Artin) Observe that the restriction of the form to a subspace W is non-degenerate if and only if $W \cap W^\perp = 0$. (Theorem 8.4.5 of Artin) Show that this happens if and only if $W \oplus W^\perp = V$. The resulting map $V \mapsto W$ (given by $v \mapsto w$ where $v = w + u$ uniquely with $w \in W$ and $u \in W^\perp$) is called the *orthogonal projection to* W.

   Solution: If $W \oplus W^\perp = V$, then of course $W \cap W^\perp = 0$. Now let us suppose that $W \cap W^\perp = 0$. We know that $\dim W^\perp \geq \dim V - \dim W$ (item (3) above), so that $\dim(W + W^\perp) = \dim W + \dim W^\perp - \dim(W \cap W^\perp) \geq n$, which means that $W + W^\perp = V$. □

7. (Theorem 8.4.5 of Artin) If the form is non-degenerate on V and restricts to a non-degenerate form on W, then its restriction to $W^\perp$ is also non-degenerate.

   Solution: To show that the form restricts non-degenerately to $W^\perp$ is equivalent to showing that $W^\perp \cap (W^\perp)^\perp = 0$ (item (6) above). But, since $V = W \oplus W^\perp$ (because the form restricts non-degenerately to W; see item (6) again), any vector in $W^\perp \cap (W^\perp)^\perp$ would be in $V^\perp$. But $V^\perp = 0$ since the form is non-degenerate on V. □

Projection Formula

8. (Projection Formula; Theorem 8.4.11 of Artin) Let $\langle\ ,\ \rangle$ be a symmetric form on a finite dimensional vector space V. Let W be a subspace the restriction to which of the form is non-degenerate. Let $\{w_1, \ldots, w_k\}$ be an orthogonal basis for W. Then the orthogonal projection p to W—see item (6) above—is given by the following formula:
$$p(v) = \sum_{i=1}^{k} \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle} w_i$$
Solution: The given formula defines a linear transformation on V that is identity on W and zero on $W^\perp$. $\square$

9. Consider the Euclidean space $\mathbb{R}^n$ with its standard inner product. Let A be an $n \times m$ matrix whose columns are linearly independent. Suppose W is the column space of A (that is, the subspace of $\mathbb{R}^n$ spanned by the columns of A). Then:

   (a) The $m \times m$ matrix A^tA invertible.

   (b) The matrix (with respect to the standard basis of $\mathbb{R}^n$) of the orthogonal projection of to W (considered as a linear operator on $\mathbb{R}^n$) is given by the following formula: $A(A^tA)^{-1}A^t$.

   Solution:

   (a) $A^tAx = 0$ implies $x^tA^tAx = 0$ implies $Ax = 0$ implies $x = 0$ (since the columns of A are linearly independent). $\square$

   (b) Let M be the suggested matrix. It is readily checked that $Mw = w$ for w in W (the column space of A) and $Mv = 0$ for $v \in W^\perp$.

10. On the space of $m \times n$ real matrices, consider the bilinear form $\langle A, B\rangle := \text{trace}(A^tB)$. Check that this is symmetric and positive definite.

    Solution: The form is symmetric because, being the transposes of each other, the matrices A^tB and B^tA have the same trace. If v_i and w_j are the i^{th} and j^{th} columns of A and B respectively (thought of as vectors in $\mathbb{R}^n$), then the entry in position (i, j) of A^tB is the usual dot product $v_i \cdot w_j$ of the vectors v_i and w_j. Thus the trace of A^tA equals $\sum_{i=1}^{n} v_i \cdot v_i$. Since $v \cdot v \geq 0$ for any v in $\mathbb{R}^n$ and $v \cdot v = 0$ only if $v = 0$ (by the positive definiteness of the usual dot product in $\mathbb{R}^n$), it follows that $\text{trace}(A^tA) \geq 0$ and that $\text{trace}(A^tA) = 0$ only if $A = 0$. $\square$

11. Let M be the space of all $n \times n$ real matrices and S the subspace of symmetric matrices.

    (a) Describe $S^\perp$, with respect to the symmetric positive definite form $\langle A, B\rangle = \text{trace}(A^tB)$ on M.

    (b) Give an explicit formula for the orthogonal projection to S.

    Solution:

    (a) The given form being positive definite, its restriction to S is also positive definite and therefore non-degenerate. Thus, by item (6) above, we have $M = S \oplus S^\perp$. We claim that $S^\perp$ is the space Z of skew-symmetric matrices. If A is symmetric and B is skew-symmetric, then, on the one hand, $\langle A, B\rangle = \text{trace}(A^tB) = \text{trace}(AB)$ and, on the other, it equals $\langle B, A\rangle = \text{trace}(B^tA) = -\text{trace}(BA) = -\text{trace}(AB)$, so $\langle A, B\rangle = 0$. This shows that Z is contained in $S^\perp$. Let us now calculate the dimension of $S^\perp$: it equals $\dim M - \dim S = n^2 - (n(n+1)/2) = n(n-1)/2$, which is the dimension of Z. Since $Z \subseteq S^\perp$ and $\dim Z = \dim S^\perp$, it follows that $Z = S^\perp$.

    (b) The orthogonal projection to S is given by the formula $A \mapsto (A + A^t)/2$. Indeed, this map is linear, identity on the space S of symmetric matrices, and vanishes on the space $S^\perp$ of skew-symmetric matrices. $\square$

Structure Theorem for Symmetric and Hermitian Forms (Sylvester)

Let V be either one of these two: an n-dimensional real vector space with a fixed symmetric form; or an n-dimensional complex vector space with a fixed Hermitian form.

12. (Lemma 8.4.9 of Artin) Suppose that our form $\langle\ ,\ \rangle$ is not identically zero. Then there exists a vector v in V such that $\langle v, v\rangle \neq 0$.

    Solution: Since the form is not identically zero, there exist vectors v and w in V such that $\langle v, w\rangle \neq 0$. If either $\langle v, v\rangle \neq 0$ or $\langle w, w\rangle \neq 0$, we are done. If not, $\langle v+w, v+w\rangle = 2\langle v, w\rangle \neq 0$, and we are done. □

13. (Theorem 8.4.10 of Artin) There exists a basis of V with respect to which the matrix of the form is diagonal. Such a basis is called an *orthogonal basis*.

    Solution: We proceed by induction on the dimension n of V. The assertion being obvious for $n = 1$, let us assume $n \geq 2$. If the form is identically zero, then its matrix is zero (with respect to any basis), and we are done. So, let us assume that it is not zero. By item (h) above, we can find a vector v in V such that $\langle v, v\rangle \neq 0$.

    Let W be the subspace spanned by the vector v. Clearly, the form restricted to W is non-degenerate. And so (see item (6) above), $V = W \oplus W^\perp$. Consider the form on $W^\perp$ obtained by restriction. By induction, we can get an orthogonal basis for $W^\perp$ (with respect to this restricted form). This basis augmented by v is an orthogonal basis for V. □

14. (**Sylvester's Structure Theorem**: Existence) There exist non-negative integers p, m, and z such that $p+m+z = n := \dim V$ and a basis of V with respect to which the matrix of the form is the "direct sum" of the three matrices $I_{p\times p}$, $-I_{m\times m}$, $0_{z\times z}$.

    Solution: By item (9) above, there exists an orthogonal basis, say $\{v_i\}_{i=1}^n$, for V. Fix such a basis. We scale each of the basis vectors as follows: if $\langle v_i, v_i\rangle \neq 0$, replace v_i by $v_i/\sqrt{|\langle v_i, v_i\rangle|}$. Thus we may assume that each $\langle v_i, v_i\rangle$ is either 0, 1, or -1. A rearrangement of the v_i has the desired properties. □

15. Let p, m, and z be as in the previous item. Then the form on V is non-degenerate if and only if $z = 0$; it is positive semi-definite if and only if $m = 0$; it is negative semi-definite if and only if $p = 0$; positive definite if and only if $m + z = 0$; negative definite if and only if $p + z = 0$.

16. (**Sylvester's Structure Theorem**: Uniqueness) The integers p, m, and z in the previous item depend only upon the form and not on the choice of basis.

    Solution: Towards the proof of uniqueness, fix an orthogonal basis as above and let: P be the subspace spanned by the first p basis vectors; M the subspace spanned by the next m of them; and Z the subspace spanned by the last z of them. We claim that Z is the null space of the form: evidently, Z belongs to the null space; and, conversely, if v is a vector in $V \setminus Z$ belonging to the null space, writing v as a linear combination $\sum_{i=1}^n a_i v_i$, we see that if any of the $a_j \neq 0$ for $1 \leq j \leq p+m$, then $\langle v, v_j\rangle \neq 0$, which proves that the null space is contained in Z. Since z is the dimension of the null space Z, it is independent of the choice of the basis.

    To prove that p depends only on the form and not on the choice of the basis, we **claim** that p is the maximal dimension of a subspace to which the form restricts positive definitely. On the one hand, the restriction of the form to P is evidently positive definite. On the other, suppose W is a subspace that is of dimension of more than p. Then W intersects $M+Z$ non-trivially. But the form restricts negative semi-definitely to $M + Z$. Choosing $0 \neq w \in W \setminus (M+Z)$, we conclude that $\langle w, w\rangle \leq 0$, which shows that the restriction of the form to W cannot be positive definite. This proves our claim. It follows that p is intrinsic to the form and does not depend on the choice of basis. Finally, since $p+m+z = n = \dim V$, it follows that m too is independent of the chosen basis. □

17. A symmetric $n \times n$ real matrix A is called *positive definite* if the symmetric form $\langle x, y\rangle = x^t A y$ on $\mathbb{R}^n$ is positive definite. Show that the following are equivalent for an $n \times n$ real matrix:

    (a) A is symmetric and positive definite;

    (b) $A = B^t B$ for some invertible $n \times n$ real matrix B;

    (c) A is symmetric and $\det A_i > 0$ for $1 \leq i \leq n$, where A_i denotes the sub-matrix of A formed by the first i rows and columns.

Solution: First suppose that A is symmetric and positive definite. Then by Sylvester's structure theorem there exists an orthonormal basis for the form $\langle x, y\rangle = x^t Ay$. Letting P be the invertible matrix whose columns are such an orthonormal basis, we have $P^t AP = I$. Putting $B = P^{-1}$, we get $A = B^t B$. This shows that (a)⇒(b).

Let us now show that (b)⇒(c). Evidently $A = B^t B$ is symmetric, and we have $\det A = \det(B^t B) = (\det B)^2$. So $\det A > 0$ (since B is invertible). This proves (c) for $i = n$. We have already shown that A is positive definite, that is to say, the form $\langle x, y\rangle := x^t Ay$ is positive definite on $\mathbb{R}^n$. The restriction of this form to any subspace is also positive definite. The restriction to the subspace spanned by the first i standard basis vectors has matrix A_i with respect to those basis vectors. Thus $\det A_i > 0$ (by invoking the case $i = n$ that we have already proved).

Finally, to prove (c)⇒(a), we proceed by induction on n. The base case $n = 1$ is clear. Proceeding to the induction step, we may, by the induction hypothesis, assume that $A' := A_{n-1}$ is positive definite. Let us denote by b the $(n-1)\times 1$ matrix obtained by looking at the last column of A and leaving out the its last element a_{nn}. Since A' is invertible, we have $A'c = b$ for some c in $\mathbb{R}^{n-1}$. We have

$$A = \begin{pmatrix} A' & b \\ b^t & a_{nn} \end{pmatrix} = \begin{pmatrix} A' & A'c \\ c^t A'^t & a_{nn} \end{pmatrix}$$

For x in $\mathbb{R}^n$ let us write x' for the top $(n-1)\times 1$ sub-matrix of x. With this understanding, let us compute $\langle x, y\rangle = x^t Ay$ for x, y in $\mathbb{R}^n$. We get $\langle x, y\rangle = x'^t A'y' + c^t A'^t y' x_n + x'^t A'cx_n + x_n^2 a_{nn}$. Putting $y = x$ and rewriting, we get:

$$\langle x, x\rangle = (x'^t + x_n c^t)A'(x' + x_n c) + (x_n)^2\left(a_{nn} - c^t A'c\right)$$

Given the positive definiteness of A', it suffices, for the positive definiteness of A, to prove that $a_{nn} - c^t A'c > 0$. But we have

$$\begin{aligned} \det A = \det\begin{pmatrix} A' & A'c \\ c^t A'^t & a_{nn} \end{pmatrix} &= \det\begin{pmatrix} A' & A'c - A'c \\ c^t A'^t & a_{nn} - c^t A'c \end{pmatrix} \\ = \det\begin{pmatrix} A' & 0 \\ c^t A'^t & a_{nn} - c^t A'c \end{pmatrix} &= \det A' \cdot (a_{nn} - c'^t A'c) \end{aligned}$$

Since $\det A$ and $\det A'$ are positive by hypothesis, it follows that $a_{nn} - c'^t A'c > 0$, and we are done. □

18. A Hermitian $n\times n$ complex matrix A is called *positive definite* if the Hermitian form $\langle x, y\rangle = x^\star Ay$ on $\mathbb{C}^n$ is positive definite. Show that the following are equivalent for an $n\times n$ complex matrix:

    (a) A is Hermitian and positive definite;

    (b) $A = B^\star B$ for some invertible $n\times n$ complex matrix B;

    (c) A is Hermitian and $\det A_i > 0$ for $1 \le i \le n$, where A_i denotes the sub-matrix of A formed by the first i rows and columns. (Note that A_i is Hermitian and therefore $\det A_i$ is a real number.)

    Solution: Entirely analogous to the proof in item (16). □

Euclidean and Hermitian Spaces

A *Euclidean space* is a finite dimensional real vector space equipped with a symmetric positive definite bilinear form. A *Hermitian space* is a finite dimensional complex vector space equipped with a Hermitian positive definite (sesqui-linear) form. Let V be a fixed Euclidean or Hermitian space. Let $\langle\ ,\ \rangle$ denote the fixed form on V and put $n := \dim V$. Applying Sylvester's Theorem (see the previous section), there exists a basis $\{v_1, \ldots, v_n\}$ of V such that $\langle v_i, v_j\rangle = \delta_{ij}$. Such a basis of V is called *orthonormal*.

The Adjoint

Let T be a linear transformation from V to itself. Given $w \in V$, consider the linear functional $v \mapsto \langle Tv, w\rangle$ on V. Since $\langle\ ,\ \rangle$ is non-degenerate, there is an element y of V such that $\langle Tv, w\rangle = \langle v, y\rangle$. We denote y by $T^\star w$. The association $w \mapsto T^\star w$ defines a linear operator on V. It is defined by the requirement that $\langle Tv, w\rangle = \langle v, T^\star w\rangle$ for all v and w in V.

1. Let M be the matrix of T with respect to an orthonormal basis. Then the matrix of the $T^\star$ with respect to the same basis is $M^\star$, the conjugate transpose of M.

   Solution: Indeed, denoting by $\{v_1, \ldots, v_n\}$ the orthonormal basis of V, we have $M_{ij} = \langle Tv_j, v_i\rangle = \langle v_j, T^\star v_i\rangle = \overline{\langle T^\star v_i, v_j\rangle} = M^\star_{ji}$. □

2. T is called *symmetric* (in the real case) or *Hermitian* (in the complex case) if $T = T^\star$. The following conditions are equivalent for T:
   (a) T is symmetric (respectively Hermitian).
   (b) $\langle Tv, w\rangle = \langle v, T^\star w\rangle$ for all v, w in V.
   (c) The matrix M of T with respect to any orthonormal basis is symmetric (respectively, Hermitian).
   (d) The matrix M of T with respect to some orthonormal basis is symmetric (respectively, Hermitian).

3. T is called *orthogonal* (in the real case) or *unitary* (in the complex case) if $TT^\star = T^\star T = I$ (that is, T and $T^\star$ are inverses of each other). The following conditions are equivalent for T:
   (a) T is orthogonal (respectively, unitary).
   (b) $\langle Tv, Tw\rangle = \langle v, w\rangle$ for all v, w in V.
   (c) The matrix M of T with respect to some any basis is orthogonal (respectively, unitary).
   (d) The matrix M of T with respect to some orthonormal basis is orthogonal (respectively, unitary).

4. T is called *normal* if $TT^\star = T^\star T$, that is T commutes with $T^\star$. The following conditions are equivalent for T:
   (a) T is orthogonal (respectively, unitary).
   (b) $\langle Tv, Tw\rangle = \langle T^\star v, T^\star w\rangle$ for all v, w in V.
   (c) The matrix M of T with respect to some any basis is orthogonal (respectively, unitary).
   (d) The matrix M of T with respect to some orthonormal basis is orthogonal (respectively, unitary).

Eigenvalues

The eigenvalues of a Hermitian operator T are all real. Indeed, if $Tv = \lambda v$ (with λ being a complex number and $0 \neq v$), then $\lambda\langle v, v\rangle = \langle Tv, v\rangle = \langle v, T^\star v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle$, and so $\lambda = \overline{\lambda}$ (note $\langle v, v\rangle \neq 0$ because of the positive definiteness of the form).

If A is a $n \times n$ Hermitian matrix, it can be thought of as being the matrix of a Hermitian operator on $\mathbb{C}^n$ with respect to the standard basis of $\mathbb{C}^n$. Since all the eigenvalues of the operator are real, so are those of A. In particular, the eigenvalues of a real symmetric matrix are real.

The eigenvalues of a unitary operator lie on the unit circle. Indeed, if $Tv = \lambda v$ (with λ being a complex number and $0 \neq v$), then $\lambda\langle v, v\rangle = \langle Tv, v\rangle = \langle v, T^\star v\rangle = \langle v, T^{-1}v\rangle = \left\langle v, \frac{1}{\lambda}v\right\rangle = \overline{\frac{1}{\lambda}}\langle v, v\rangle$, and so $\lambda\overline{\lambda} = 1$ (note $\langle v, v\rangle \neq 0$ because of the positive definiteness of the form).

Spectral Theorem

In these exercises we prove the **Spectral Theorem**. Let V be a finite dimensional *Hermitian space*: this means that V is a finite dimensional complex vector space along with a fixed Hermitian positive definite form $\langle\ ,\ \rangle$. Let T be a linear operator on V. Recall that we have defined the adjoint $T^\star$ (as a linear operator on V) using the form: it is uniquely determined by the requirement that $\langle Tv, w\rangle = \langle v, T^\star w\rangle$ for all v, w in V; alternatively (but equivalently), it is defined as that operator whose matrix with respect to some orthonormal basis—such a basis exists by Sylvester's structure theorem—is the conjugate transpose of the matrix of T with respect to that basis.

1. Suppose that W is a T-invariant subspace of V. Then $W^\perp$ is $T^\star$-invariant.

   Solution: For v in $W^\perp$ and w in W, we have $\langle T^\star v, w\rangle = \langle v, Tw\rangle = 0$ (since $Tw \in W$). This shows that $T^\star v$ belongs to $W^\perp$. □

2. Suppose that T is *normal* (that is, T and $T^\star$ commute). Then:

   (a) If $Tv = 0$ for some v in V. Then $T^\star v = 0$.

   (b) If $Tv = \lambda v$ then $T^\star v = \overline{\lambda} v$, and $\langle v\rangle^\perp$ is T-invariant.

   Solution:

   (a) We have $0 = \langle Tv, Tv\rangle = \langle T^\star Tv, v\rangle = \langle TT^\star v, v\rangle = \langle T^\star v, T^\star v\rangle$, which proves that $T^\star v = 0$. □

   (b) Put $S := T - \lambda I$. Then $SS^\star = (T - \lambda I)(T^\star - \overline{\lambda} I) = (T^\star - \overline{\lambda} I)(T - \lambda I) = S^\star S$, so S is normal. Since $Sv = 0$, it follows from the previous item that $S^\star v = 0$, or in other words that $T^\star v = \overline{\lambda} v$. For the second assertion, since $\langle v\rangle$ is $T^\star$-invariant, it follows from item (1) above that $\langle v\rangle^\perp$ is T-invariant. □

3. Let W be a subspace that is T-invariant and $T^\star$-invariant. Let S denote the restriction of T to W. Then $S^\star$ (the adjoint of S calculated with respect to the Hermitian space W) equals the restriction of $T^\star$ to W.

   Solution: For v and w in W, we have $\langle Sv, w\rangle = \langle Tv, w\rangle = \langle v, T^\star w\rangle$. But $\langle Sv, w\rangle = \langle v, S^\star w\rangle$ (by definition of $S^\star w$). Since $\langle v, T^\star w\rangle = \langle v, S^\star w\rangle$ and $T^\star w$ belongs to W, for all v and w in W, we conclude that $S^\star$ is the restriction of $T^\star$ to W. □

4. Prove the **Spectral Theorem** (for Symmetric and Hermitian Operators): Let T be a symmetric (respectively Hermitian) linear operator on a finite dimensional Euclidean (respectively Hermitian) space V. Then there exists an orthonormal basis of V consisting of eigenvectors for T.

   Solution: Proceed by induction on the dimension n of V. The case $n = 1$ being obvious—choose any vector of unit norm as basis—let us assume that $n \geq 2$ and that the assertion holds whenever V has dimension less than n. Let v in V be an eigenvector for T (in the real Symmetric case, we have seen already that all the eigenvalues are real, and so given any eigenvalue we can find an eigenvector in V with that eigenvalue). Scaling v as needed, we may assume that $\langle v, v\rangle = 1$.

   From item (1) above, it follows that $\langle v\rangle^\perp$ is invariant under $T^\star = T$. From item (3), it follows that the adjoint of the restriction—call it S—of T to $\langle v\rangle^\perp$ is the restriction of $T^\star = T$ to it, and so $S = S^\star$. It follows from the induction hypothesis that there exists an orthonormal basis of $\langle v\rangle^\perp$ consisting of eigenvectors of S. These basis elements for $\langle v\rangle^\perp$ along with v clearly form an orthonormal basis of eigenvectors for V. □

5. Prove the **Spectral Theorem** (for Normal Operators): Let T be a normal linear operator on a finite dimensional Hermitian space V. Then there exists an orthonormal basis of V consisting of eigenvectors for T.

   Solution: Proceed by induction on the dimension n of V. The case $n = 1$ being obvious—choose any vector of unit norm as basis—let us assume that $n \geq 2$ and that the assertion holds whenever V has dimension less than n. Let v be an eigenvector for T. Scaling v as needed, we may assume that $\langle v, v\rangle = 1$.

   From item (2b) above, it follows that $\langle v\rangle^\perp$ is invariant under T; and from item (1) that it is $T^\star$-invariant. By item (3), it follows that the adjoint of the restriction—call it S—of T to $\langle v\rangle^\perp$ is the restriction of $T^\star$ to it, and so S commutes with $S^\star$. Since S is normal, it follows from the induction hypothesis that there exists an orthonormal basis of $\langle v\rangle^\perp$ consisting of eigenvectors of S. These basis elements for $\langle v\rangle^\perp$ along with v clearly form an orthonormal basis of eigenvectors for V. □

6. Let T be a normal operator on a finite dimensional Hermitian space V. Show that any T-invariant subspace of V is also $T^\star$-invariant. (Note: Item (2b) above shows this in the case when the subspace is 1-dimensional.)

   SOLUTION: We give two different proofs, the first one using the Spectral Theorem, the second one more direct. For either proof, let W be a fixed T-invariant subspace of V.

   (a) (First proof, using the Spectral Theorem) Since T is diagonalizable (by the Spectral Theorem), its restriction to W is too. Thus W has a basis of eigenvectors for T. Each of these is invariant under $T^\star$ by item (2b) above, and so W too is $T^\star$-invariant.

   (b) (A more direct proof) Choose an orthonormal basis for W and extend this (e.g., by Gram-Schmidt) to an orthonormal basis for V. In terms of such a basis for V, the matrices of T and $T^\star$ look like:

   $$T = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \qquad T^\star = \begin{pmatrix} A^\star & 0 \\ B^\star & D^\star \end{pmatrix}$$

   and so:

   $$TT^\star = \begin{pmatrix} AA^\star + BB^\star & A^\star B \\ DB^\star & DD^\star \end{pmatrix} \qquad T^\star T = \begin{pmatrix} A^\star A & A^\star B \\ B^\star A & B^\star B + D^\star D \end{pmatrix}$$

   Since T is normal, we have $TT^\star = T^\star T$ which means in particular that $AA^\star + BB^\star = A^\star A$. Taking traces we get $\text{trace}(BB^\star) = 0$ (because $\text{trace}(AA^\star) = \text{trace}(A^\star A)$). But this means $B = 0$ (by the Hermitian analogue of item (10) on Page 4). Thus $B^\star = 0$, which means that W is $T^\star$-invariant.

Matrix Versions of the Spectral Theorem

1. Let A be a $n \times n$ Hermitian, unitary, or just a normal matrix: recall that an $n \times n$ complex matrix A is called *Hermitian* if $A = A^\star$ (where $A^\star$ denotes the conjugate transpose of A); unitary if $A^\star$ is the inverse of A (that is $AA^\star = A^\star A = I$); and *normal* if A and $A^\star$ commute (that is, $AA^\star = A^\star A$). Then there exists a unitary $n \times n$ matrix P such that $P^\star AP = P^{-1}AP$ is diagonal. The diagonal entries are real if A is Hermitian, and lie on the unit circle in the complex plane if A is unitary. In particular, there is a diagonal matrix in every conjugacy class of unitary matrices.

2. Let A be a $n \times n$ real symmetric matrix. Then there exists an $n \times n$ real orthogonal matrix P such that $P^t AP = P^{-1}AP$ is diagonal. In particular there is a diagonal matrix in every conjugacy class of real orthogonal matrices.

PROOF:

1. Consider the standard Hermitian space $\mathbb{C}^n$ with its standard Hermitian inner product. Let T be the operator on $\mathbb{C}^n$ whose matrix with respect to the standard basis is the given matrix A. Since the standard basis is orthonormal with respect to the standard inner product, it follows that T is Hermitian (that is, $T^\star = T$), unitary (that is, $T^\star T = TT^\star = I$), or normal (that is, T commutes with $T^\star$) accordingly as A.

   By the Spectral Theorem, there exists an orthonormal basis—say $v_1, \ldots, v_n$—of $\mathbb{C}^n$ consisting of eigenvectors for T. Put $P := [v_1, \ldots, v_n]$ (the i^{th} column of P consists of the co-ordinates of v_i). Then $P^\star = [v_1, \ldots, v_n]^\star$ (the i^{th} row of $P^\star$ consists of the conjugate-transposes of the co-ordinates of v_i). The fact that the basis $v_1, \ldots, v_n$ is orthonormal translates to $P^\star P = I$, so P is unitary. The fact that $v_1, \ldots, v_n$ are eigenvectors for T translates to $AP = PD$ where D is the diagonal matrix with the entry at position (i, i) being the eigenvalue of v_i. Left multiplying by $P^\star$ we get $P^\star AP = P^{-1}AP = D$.

   As for the statement about eigenvalues of A (in the case when it is Hermitian or unitary), we have seen earlier that the eigenvalues of a Hermitian operator (on a finite dimensional Hermitian space) are real and that of a unitary operator are unit complex numbers. We can adapt those arguments for matrices. Alternatively, we can think of A as the matrix with respect to the standard basis of $\mathbb{C}^n$ of a Hermitian or unitary operator as the case may be, and then apply the result for operators.

2. Consider the standard Euclidean space $\mathbb{R}^n$ with its standard inner product. Let T be the operator on $\mathbb{R}^n$ whose matrix with respect to the standard basis is the given matrix A. Since the standard basis is orthonormal with respect to the standard inner product, it follows that T is symmetric. We apply the Spectral Theorem to the symmetric operators and then proceed as in the proof of part (1) above. □

Skew-Symmetric Forms

In these exercises, we define **Skew-Symmetric Forms** and prove a **Structure Theorem** for them. Let V be a finite dimensional vector space over some field k and put $n := \dim V$. A bilinear form $\langle\ ,\ \rangle$ is called *Skew-Symmetric* if $\langle v, v\rangle = 0$ for all $v \in V$.

1. Supposing that our bilinear form $\langle\ ,\ \rangle$ is skew-symmetric, show that $\langle v, w\rangle = -\langle w, v\rangle$ for all v, w in V. Prove the following converse: supposing that the form satisfies $\langle v, w\rangle = -\langle w, v\rangle$ for all v, w in V, show that it is skew-symmetric, provided that the characteristic of k is not 2. This explains why we have preferred to define a skew-symmetric form as we've done rather than just insist on $\langle v, w\rangle = -\langle w, v\rangle$. Our definition works—in the sense that the results below hold—without exception, even if the field happens to have characteristic 2.
2. Consider the case $k = \mathbb{R}$ and $V = \mathbb{R}^2$. Define a skew-symmetric form on $\mathbb{R}^2$ by $\langle x, y\rangle = x^t J y$, where J is the following skew-symmetric matrix:
$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
This form on $\mathbb{R}^2$ is called the *standard area form*. And $\mathbb{R}^2$ equipped with the above form is called the *standard symplectic plane*. The structure theorem proved below says that any real finite dimensional vector space equipped with a non-degenerate skew-symmetric form is isomorphic to a product of such planes.
3. We have seen earlier that the vector space of bilinear forms on V is isomorphic to the space of $n \times n$ matrices, albeit non-canonically, depending upon the choice of a basis.
   (a) Observe that the space of all skew-symmetric forms on V is a subspace of the space of all bilinear forms. On the other hand, the set of all skew-symmetric $n \times n$ matrices A is also a subspace of the space of $n \times n$ matrices. (In case the characteristic of k is 2, we demand that the diagonal entries be zero for a matrix to be considered skew-symmetric.) When the isomorphism of between bilinear forms and matrices is restricted to the subspace of skew-symmetric forms, we get an isomorphism of this subspace onto the subspace of skew-symmetric matrices. What is the common dimension of these subspaces? (Answer: $n(n-1)/2$.)
   (b) The isomorphism in the previous item can be made more explicit as follows. Think of V as being k^n (the space of $n \times 1$ matrices). Given a (skew-symmetric) matrix A, associate with it the (skew-symmetric) form on k^n given by $\langle x, y\rangle = x^t A y$.
4. Suppose that our form $\langle\ ,\ \rangle$ on V is skew-symmetric and non-degenerate. Prove that the dimension of V must be even (even when the field has characteristic 2?!).

   Solution: Let $A_{\mathcal{B}}$ be the matrix of the form with respect to some basis $\mathcal{B}$. It is both skew-symmetric and invertible. But, as an elementary calculation shows, the determinant of an $n \times n$ skew-symmetric matrix vanishes if n is odd, provided the characteristic of k is not 2. How does one manage characteristic 2? $\square$
5. Prove the following **Structure Theorem**. Let V be a finite dimensional vector space with a non-degenerate skew-symmetric form $\langle\ ,\ \rangle$. Then the dimension of V is even and there exists a basis of V with respect to which the matrix of the form is of block diagonal form, with each block being the following matrix J:
$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
Solution: We proceed by induction on $n := \dim V$. In the case n is zero, the assertion is obvious (we choose the only possible basis, namely the empty set). Let's therefore assume that $n \geq 1$ and move to the induction step of the proof. Let v be any non-zero element of V. Since the form is non-degenerate, there exists w in V such that $\langle v, w\rangle \neq 0$ (note that w cannot be a multiple of v, because the form is skew-symmetric, so that the span of v and w is a two dimensional subspace). Scaling v (or w for that matter), we may assume that $\langle v, w\rangle = -\langle w, v\rangle = 1$. Let W be the span of v and w. The matrix of the form restricted to W is J with respect to the basis v, w, so the form restricted to W is non-degenerate (by item (5) on Page 3 above, since $\det J \neq 0$). By item (6) on Page 3, $V = W \oplus W^\perp$, so $\dim W^\perp = n - 2$; and by item (7) on Page 3, the form restricted to $W^\perp$ is non-degenerate.

   Since $\dim W^\perp < \dim V$, we can apply the induction hypothesis to get a basis for $W^\perp$ with the desired property. That basis augmented with v, w is a basis of V with the desired property. $\square$

On the Space of Real Symmetric Positive Definite Matrices

Let A be an $n \times n$ real matrix, and denote by $\langle x, y \rangle_A := x^t A y$ the form on $\mathbb{R}^n$ defined by A.

- The matrix A is symmetric $\Leftrightarrow$ the form $\langle\ ,\ \rangle_A$ is symmetric.
- A symmetric matrix A is said to be *positive definite* if the symmetric form $\langle\ ,\ \rangle_A$ is positive definite.

Now assume that A is symmetric and positive definite. The space $\mathbb{R}^n$ with the form $\langle\ ,\ \rangle_A$ on it is a Euclidean Space (as defined on Page 7). By Sylvester's Structure Theorem (Page 5), it admits an orthonormal basis. If we denote by B the $n \times n$ matrix whose columns form an orthonormal basis, we have $B^t A B = I$ or $A = C^t C$ where $C := B^{-1}$.

Let $\mathcal{S}$ be the set of all symmetric and positive definite $n \times n$ real matrices. There is an action of the group $G = GL_n(\mathbb{R})$ of real $n \times n$ invertible matrices on $\mathcal{S}$ as follows: $g \cdot A := gAg^t$. The above observation that each element A of $\mathcal{S}$ is of the form $C^t C$ for some invertible matrix C can be expressed in terms of this action as follows: **the action of G on $\mathcal{S}$ is transitive**. Thus we can think of $\mathcal{S}$ as a homogeneous space for this action of G. The stabiliser of the identity matrix (as an element of $\mathcal{S}$) being the subgroup $O_n(\mathbb{R})$ of orthogonal matrices, we have:

$$\mathcal{S} = GL_n(\mathbb{R})/O_n(\mathbb{R})$$

In the sequel, $\mathcal{U}$ denotes the set of all real $n \times n$ invertible upper triangular matrices with positive diagonal entries.

1. Given A in $\mathcal{S}$, there is a unique matrix D in $\mathcal{U}$ such that $A = D^t D$.

   Solution: Both the existence and uniqueness of D can be seen by elementary arguments. To begin with d_{11} is the (unique) positive square root of a_{11} (note that all principal minors, and in particular all the diagonal entries, of A are positive). Then, d_{1j} for $j > 1$ are uniquely determined by the requirement that $d_{1j} d_{11} = a_{1j}$. The element d_{22} is uniquely determined by the conditions that it be positive and satisfy $d_{12}^2 + d_{22}^2 = a_{22}$: note that $a_{22} - d_{12}^2 = (a_{11}a_{22} - a_{12}^2)/a_{11} > 0$, so the conditions can be met. The numbers d_{2j} for $j > 2$ can now be uniquely determined by the condition that $d_{12}d_{1j} + d_{22}d_{2j} = a_{2j}$. And so on.

   But we can also obtain this as an application of the Gram Schmidt process (see below). $\square$

2. (Gram Schmidt) Given a symmetric positive definite form on $\mathbb{R}^n$ and any invertible matrix M, there exists a unique element E of $\mathcal{U}$ such that the columns of the matrix ME form a basis of $\mathbb{R}^n$ orthonormal with respect to the form.

   Solution: This is clear from the usual Gram Schmidt process applied to the basis formed by the columns of M. $\square$

3. Prove the claim in item (1) using the Gram Schmidt process (item (2)).

   Solution: Let A be an arbitrary element of $\mathcal{S}$ and $\langle\ ,\ \rangle_A$ the corresponding symmetric positive definite form on $\mathbb{R}^n$. Invoke item (2) with M as the identity matrix I, to get E in $\mathcal{U}$ whose columns are orthonormal with respect to $\langle\ ,\ \rangle_A$. This means that $E^t A E = I$, or $A = D^t D$, where $D = E^{-1}$. The uniqueness of D follows from the uniqueness of E. $\square$

4. Prove the lower triangular analogues of items (1) and (2). More precisely, letting $\mathcal{L}$ be the set of all lower triangular invertible $n \times n$ matrices with positive diagonal entries, show that

   (a) given a positive definite form on $\mathbb{R}^n$ and an invertible $n \times n$ matrix M, there exists a unique element L of $\mathcal{L}$ such that the columns of the matrix ML are orthonormal with respect to the form. (Hint: This is again the usual Gram Schmidt, this time applied to the columns of M in reverse order.)

   (b) for any A in $\mathcal{S}$ there exists a unique element L in $\mathcal{L}$ such that $L^t L = A$.

Negative Definite Matrices

5. A symmetric $n \times n$ real matrix A is called *negative definite* if the symmetric form $\langle x, y \rangle := x^t A y$ on $\mathbb{R}^n$ is negative definite. This property of a symmetric matrix is equivalent to any of these: (a) $-A$ is positive definite; (b) there exists an invertible matrix B such that $A = -B^t B$; (c) there exists a unique element U of $\mathcal{U}$ such that $A = -U^t U$; (d) there exists a unique element of L of $\mathcal{L}$ such that $A = -L^t L$; (e) $(-1)^i \det A_i > 0$ where A_i is the sub-matrix of A consisting of its first i rows and columns.