

Math 3410 Assignment #5

University of Lethbridge, Spring 2015

Sean Fitzpatrick

April 7, 2015

Due date: Thursday, April 2nd, by 5 pm.

Please provide solutions to the problems below, using the usual guidelines. In particular, you should include a cover page and a reasonable good copy. Your cover page should reference any resources used to complete the assignment, including websites and people. The rules about how you're not supposed to cheat on your assignment remain the same as for previous assignments. In particular, you must **cite all sources**.

Assigned problems

Before you begin, type the following phrases into Google¹, and read the first one or two websites that come up: “orthogonal diagonalization”, “positive-definite matrix”, “principal axes theorem”, “self-adjoint operator”.

For the purposes of this assignment, \mathbb{R}^n (or \mathbb{C}^n) will denote the vector space of $n \times 1$ column vectors, and we will work with $n \times n$ matrices rather than operators.

1. Recall that an $n \times n$ (real) matrix is *symmetric* if $A^T = A$. Suppose that A is symmetric, and λ_1 and λ_2 are distinct eigenvalues of A with corresponding eigenvectors v_1 and v_2 . Prove that v_1 and v_2 are orthogonal with respect to the standard dot product on \mathbb{R}^n .

Hint: First show that if A is symmetric, then $x \cdot (Ay) = (Ax) \cdot y$ for any $x, y \in \mathbb{R}^n$. (Recall that $x \cdot y = y^T x$.)

If you're working over \mathbb{R} : since A is symmetric, we have $A^T = A$, and thus

$$x \cdot (Ay) = (Ay)^T x = (y^T A^T) x = y^T (Ax) = (Ax) \cdot y.$$

If you're working over \mathbb{C} : since A is Hermitian, we have $A^\dagger = A$, and thus

$$\langle x, Ay \rangle = (Ay)^\dagger x = (y^\dagger A^\dagger) x = y^\dagger (Ax) = \langle Ax, y \rangle.$$

In either case, we have $\langle Ax, y \rangle = \langle x, Ay \rangle$, where $\langle \cdot, \cdot \rangle$ denotes either the real or complex dot product, depending on the context.

¹Or some other search engine of your choice. I hear there's something called Bing?

Now (over \mathbb{R} or \mathbb{C}), suppose that v_1 and v_2 are eigenvectors of A corresponding to distinct eigenvalues $\lambda_1 \neq \lambda_2$. Then if A is symmetric (Hermitian) we have

$$\lambda_1 \langle v_1, v_2 \rangle = \langle \lambda_1 v_1, v_2 \rangle = \langle Av_1, v_2 \rangle = \langle v_1, Av_2 \rangle = \langle v_1, \lambda_2 v_2 \rangle = \overline{\lambda_2} \langle v_1, v_2 \rangle.$$

In the real case, we automatically have $\overline{\lambda_2} = \lambda_2$, and since $\lambda_1 \langle v_1, v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle$, we have

$$(\lambda_1 - \lambda_2) \langle v_1, v_2 \rangle = 0,$$

and since $\lambda_1 - \lambda_2 \neq 0$ by assumption, we must have $\langle v_1, v_2 \rangle = 0$, and thus v_1 and v_2 are orthogonal, as required.

In the complex case, the same proof applies, but we first need the following result:

Lemma (And bonus question 6(a)): If A is a Hermitian matrix, then the eigenvalues of A are real.

Proof. Suppose that A is Hermitian, so that $\langle Av, w \rangle = \langle v, Aw \rangle$ for any $v, w \in V$, as shown above, and that λ is an eigenvalue of A . Let v be a corresponding eigenvector, so that $Av = \lambda v$. By multiplying by $1/\|v\|$ if necessary, we may assume that $\|v\| = 1$. Then we have

$$\lambda = \lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle Av, v \rangle = \langle v, Av \rangle = \langle v, \lambda v \rangle = \overline{\lambda} \langle v, v \rangle = \overline{\lambda}.$$

Since $\overline{\lambda} = \lambda$, we must have $\lambda \in \mathbb{R}$. □

2. **(Do not submit)** Recall that an $n \times n$ matrix P is *orthogonal* if $P^T = P^{-1}$. Prove that the columns of P form an orthonormal basis of \mathbb{R}^n .

Proof. Let C_1, \dots, C_n denote the columns of P , and let R_1, \dots, R_n denote the rows of P^T . Then by definition of transpose, we have $R_i = C_i^T$ for all $i = 1, \dots, n$. If $P^T = P^{-1}$, then we must have $P^T P = I_n$. But the (i, j) -entry of $P^T P$ is equal to

$$R_i C_j = C_i^T C_j = C_i \cdot C_j,$$

and the (i, j) -entry of I_n is $\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$, and thus we have $C_i \cdot C_j = \delta_{ij}$ for all $i, j = 1, \dots, n$, which is the definition of an orthonormal basis. □

Note: The same proof works in the complex case, with A^T replaced by A^\dagger . A complex matrix A such that $A^\dagger = A^{-1}$ is called a *unitary* matrix. As above, the i^{th} row of A^\dagger is $C_i^\dagger = \overline{C_i}^T$, and we get

$$\delta_{ij} = R_i C_j = C_i^\dagger C_j = \langle C_j, C_i \rangle.$$

One can prove that every symmetric matrix A can be *orthogonally diagonalized*; that is, there exists an orthogonal matrix P such that $PAP^T = D$ is diagonal. Using this result, solve the following problem:

3. (**Do not submit**) An $n \times n$ matrix A is called *positive-definite* if $X^T A X > 0$ for all non-zero $X \in \mathbb{R}^n$. Prove that a symmetric matrix A is positive-definite if and only if every eigenvalue of A is positive.

Suppose that A is symmetric and positive definite. Since A is a symmetric matrix, there exists an orthogonal matrix P such that $P^T A P = D$, where D is a diagonal matrix. Let $\lambda_1, \dots, \lambda_n$ denote the entries on the main diagonal of D (which are also the eigenvalues of A).

Let E_1, E_2, \dots, E_n denote the standard basis vectors for \mathbb{R}^n . Since A is positive definite, for each $i = 1, \dots, n$ we must have

$$\lambda_i = \lambda_i \langle E_i, E_i \rangle = E_i^T D E_i = E_i^T (P^T A P) E_i = (P E_i)^T A (P E_i) > 0,$$

since A is positive-definite. (Since P is invertible and $E_i \neq 0$, we must have $P E_i \neq 0$.)

Conversely, suppose A is symmetric and every eigenvalue of A is positive. Let $X = [x_1 \ x_2 \ \dots \ x_n]^T$ be any non-zero vector in \mathbb{R}^n , and let $Y = P^T X = [y_1 \ y_2 \ \dots \ y_n]^T$. Note that Y must also be non-zero. Then

$$X^T A X = X^T (P D P^T) X = (P^T X)^T D (P^T X) = Y^T D Y = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2 > 0,$$

so A is positive-definite.

Note: It's actually sufficient to assume that $A = P D P^T$, where P is orthogonal, since this implies that $A^T = (P D P^T)^T = (P^T)^T D^T P^T = P D P^T$, since $D^T = D$ for any diagonal matrix D .

A similar proof works in the complex case. If A is Hermitian, then there exists a (*real*, by the lemma above) diagonal matrix D and a unitary matrix U such that $A = U D U^\dagger$. (And conversely, if $A = U D U^\dagger$, then A is Hermitian.) The definition of positive-definite in the complex setting is

$$Z^\dagger A Z > 0 \text{ for all column vectors } Z \in \mathbb{C}^n,$$

and the proof above can be adapted to show that a Hermitian matrix is positive-definite if and only if its eigenvalues are all positive.

4. Let A be a symmetric, positive-definite matrix. Prove that $\langle X, Y \rangle = Y^T A X$ defines an inner product on \mathbb{R}^n .

Proof. Define a map $\langle \cdot, \cdot \rangle : \mathbb{F}^n \times \mathbb{F}^n \rightarrow \mathbb{F}$ by $\langle X, Y \rangle = Y^T A X$, where A is symmetric (Hermitian) and positive-definite. We need to verify that this map satisfies the definition of an inner product, so we verify the properties one-by-one.

- **Positivity:** Let $X \in \mathbb{F}^n$ be a nonzero vector. Then $\langle X, X \rangle = X^\dagger A X > 0$, by the definition of a positive-definite matrix.

- Definiteness: If $X = 0$, then clearly $\langle 0, 0 \rangle = 0^T A 0 = 0$. If $X^\dagger A X = 0$ for some $X \in \mathbb{F}^n$, then we must have $X = 0$, since otherwise A would not be positive-definite.
- Additivity: For any $X, Y, Z \in \mathbb{F}^n$, the properties of matrix multiplication show that

$$\langle X + Y, Z \rangle = Z^\dagger A(X + Y) = Z^\dagger A X + Z^\dagger A Y = \langle X, Z \rangle + \langle Y, Z \rangle.$$

- Homogeneity: For any $X, Y \in \mathbb{F}^n$ and $\alpha \in \mathbb{F}$, using the properties of matrix multiplication we have

$$\langle \alpha X, Y \rangle = Y^\dagger A(\alpha X) = \alpha(Y^\dagger A X) = \alpha \langle X, Y \rangle.$$

- Conjugate symmetry: Since A is symmetric/Hermitian, for any $X, Y \in \mathbb{F}^n$ we have

$$\langle Y, X \rangle = X^\dagger (A Y) (X^\dagger A) Y = (A^\dagger X)^\dagger Y = (A X)^\dagger Y = (Y^\dagger (A X))^\dagger = \langle X, Y \rangle^\dagger = \overline{\langle X, Y \rangle},$$

since if λ is a scalar, then $\lambda^\dagger = \bar{\lambda}$.

Since all five properties of an inner product are satisfied, we've completed our proof. \square

5. A *quadratic form* on \mathbb{R}^n is an expression of the form

$$q(x_1, \dots, x_n) = \sum_{i \leq j} a_{ij} x_i x_j.$$

For example, $q(x, y) = 2x^2 - 3xy + y^2$ and $q(x, y, z) = x^2 + xy + 2yz - xz$ are quadratic forms. There is a one-to-one correspondence between quadratic forms and symmetric matrices. If A is symmetric, we can define a quadratic form by

$$q(X) = X^T A X, \text{ where } X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

- (a) For the quadratic form $q(x, y) = 5x^2 - 4xy + 5y^2$ find a symmetric matrix A such that $q(x, y) = \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix}$.

If $A = \begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}$, then A is symmetric, and

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 5x - 2y \\ -2x + 5y \end{bmatrix} = 5x^2 - 2xy - 2yx + 5y^2 = 5x^2 - 4xy + 5y^2,$$

as required.

(b) Diagonalize the matrix A from part (a).

We know that λ is an eigenvalue of A if and only if

$$0 = \begin{vmatrix} 5 - \lambda & -2 \\ -2 & 5 - \lambda \end{vmatrix} = \lambda^2 - 10\lambda + 25 - 4 = (\lambda - 3)(\lambda - 7),$$

so the eigenvalues of A are $\lambda = 3$ and $\lambda = 7$.

For $\lambda = 3$ we have $A - 3I = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$, and we can see that a basic eigenvector for $\text{null}(A - 3I)$ is $v_1 = \begin{bmatrix} 1 & 1 \end{bmatrix}$.

For $\lambda = 7$, we have $A - 7I = \begin{bmatrix} -2 & -2 \\ -2 & -2 \end{bmatrix}$, so a basic eigenvector for $\text{null}(A - 7I)$ is $v_2 = \begin{bmatrix} 1 & -1 \end{bmatrix}$.

We note that v_1 and v_2 are orthogonal, as expected. If we let

$$P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

then P is orthogonal (as expected), and also symmetric (an added bonus). Thus, $P^T = P^{-1} = P$, and we check that

$$P^T A P = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix} = D,$$

as expected.

(c) Identify the equation $5x^2 - 4xy + 4y^2 = 1$ as a conic section, and sketch its graph.

From part (a), we know that the equation can be written as $\begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} = 1$, where A is the symmetric matrix above. From part (b), we know that we can write $A = PDP$ (we should have $A = PDP^T$ but $P^T = P$ in this case), and thus the equation is

$$\begin{aligned} \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} x & y \end{bmatrix} (PDP) \begin{bmatrix} x \\ y \end{bmatrix} = (\begin{bmatrix} x & y \end{bmatrix} P) A (P \begin{bmatrix} x \\ y \end{bmatrix}) \\ &= \begin{bmatrix} u & v \end{bmatrix} D \begin{bmatrix} u \\ v \end{bmatrix} = 3u^2 + 7v^2 = 1, \end{aligned}$$

where $\begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$. Thus, we recognize the equation as that of an ellipse, with respect to the coordinates u and v .

To sketch the graph, you need to do two things:

1. Make sure you know what the graph $3x^2 + 7y^2 = 1$ looks like.
2. Figure out how to impose the u and v axes over the x and y axes.

I'll leave you to figure out (1), perhaps by asking Wolfram Alpha. For (2), note that the u -axis corresponds to the set of points where $v = 0$, and since $v = \frac{1}{\sqrt{2}}(x - y)$, the u -axis must correspond to the line $y = x$. Similarly, the v -axis must correspond to $u = 0$, which is the line $y = -x$.

It follows that the ellipse is given by taking the ellipse $3x^2 + 7y^2 = 1$ and rotating it counter-clockwise through an angle of $\pi/4$. If you want to verify your result, you can again ask Wolfram Alpha to plot the curve given by $5x^2 - 4xy + 5y^2 = 1$. The result should look something like the sketch below:

