Linear Algebra Working Group :: Day 2

Note: All vector spaces will be finite-dimensional vector spaces over the field $\mathbb{R}$.

1 Diagonalization

Definition 1.1. An $n \times n$ matrix A is **similar** to an $n \times n$ matrix B if there exists an invertible $n \times n$ matrix P such that $A = PBP^{-1}$. More generally, let V be a finite-dimensional vector space. Linear transformations $A, B : V \to V$ are **similar** if there exists an invertible $P : V \to V$ such that $A = P \circ B \circ P^{-1}$.

Exercise 1. Show that if two matrices are similar then they have the same eigenvalues. (*Hint: Consider the characteristic polynomials.*) This of course also applies to linear transformations too. Give a counterexample to show the converse is not true.

Definition 1.2. An $n \times n$ matrix A is **diagonalizable** if it is similar to a diagonal matrix. More generally, a linear transformation $T : V \to V$ is **diagonalizable** if its matrix representation with respect to *some* basis on V is diagonalizable.

Theorem 1.3. Let A be an $n \times n$ matrix. The matrix A is diagonalizable if and only if it has n linearly independent eigenvectors.

Exercise 2. Prove Theorem Thm 1.3. (Hint: Look for a diagonalization $A = PDP^{-1}$ where D has the eigenvalues on the diagonal and P has the respective eigenvectors as columns.)

Exercise 3. Diagonalize the following matrices if possible:

$$\begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix} \qquad \begin{pmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{pmatrix}$$

Exercise 4. Show that a diagonalization is not unique.

Definition 1.4. Let A be an $n \times n$ matrix and λ an eigenvalue of A. The **algebraic multiplicity** of λ is the multiplicity of λ as a root of the characteristic polynomial $\det(A - \lambda I)$. The **geometric multiplicity** of λ is the dimension of the eigenspace E_λ of the eigenvalue λ.

Exercise 5. Suppose A is a matrix given in block form by:

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

where B and D are squares matrices. Give the eigenvalues of A, with their corresponding algebraic multiplicities, in terms of those of B and D.

Theorem 1.5. Let A be an $n \times n$ matrix and $\lambda_1, ..., \lambda_p$ its distinct eigenvalues. Let d_k be the geometric multiplicty of λ_k, and a_k the algebraic multiplicity.

1. For all $1 \leq k \leq p$, we have $d_k \leq a_k$.
2. The matrix A is diagonalizable if and only if $\sum_{k=1}^{p} d_k = n$.

3. The matrix A is diagonalizable if and only if the characteristic polynomial factors into linear factors in $\mathbb{R}$ and $\mathrm{d}_k = a_k$ for all $1 \leq k \leq p$.

4. If A is diagonalizable, the union of the bases of each eigenspace forms a basis for $\mathbb{R}^n$.

Exercise 6. Prove Theorem 1.5

Exercise 7. Determine if the matrix A is diagonalizable:

1. A is 5×5 and has two distinct eignevalues λ_1 and λ_2 with geometric multiplcities $d_1 = 3$ and $d_2 = 2$.

2. A is 4×4 and has three eigenvalues λ_1, λ_2, and λ_3, where the first two have geometric multiplicites $\mathrm{d}_1 = 1$ and $d_2 = 2$. Can A fail to be diagonalizable?

Exercise 8. Show that if an $n \times n$ matrix A has n linearly independent eigenvectors, then so does A^T.

Exercise 9. Show by a 2×2 nonzero matrix example that a matrix may be invertible, but not diagonalizable. Show by a nondiagonal 2×2 matrix that a matrix may be diagonal but not invertible.

Definition 1.6. Let A be an $n \times n$ matrix. A (real) **Schur decomposition** is a factorization of the form $A = URU^T$, where U is an orthogonal $n \times n$ matrix and R is an $n \times n$ upper triangular matrix.

Exercise 10. Let A be an $n \times n$ matrix.

1. Show that if A admits a real Schur decomposition, then A has n real eigenvalues, counting algebraic multiplicies.

2. Suppose A has n real eigenvalues $\lambda_1, ..., \lambda_n$, counting algebraic multiplicities. Let u_1 be a unit eigenvector for λ_1. Complete this to an orthonormal basis $\{u_1, ..., u_n\}$ of $\mathbb{R}^n$. Let U be the matrix with columns the vectors u_i. Show that the matrix $U^T AU$ has the following form:
$$\begin{pmatrix} \lambda_1 & * & * & * & * \\ 0 & & & & \\ \vdots & & A_1 & & \\ 0 & & & & \end{pmatrix}$$
where A_1 has eigenvalues $\lambda_2, ..., \lambda_n$. (Hint: For the last part, exercise 5 may be useful.)

3. Use part (2) to give an algorithm to obtain a real Schur decomposition when A has n real eigenvalues, counting algebraic multiplicities.

2 Symmetric Matrices, the Spectral Theorem, and Quadratic Forms

Definition 2.1. An $n \times n$ matrix A is **symmetric** or **self-adjoint** if $A^T = A$. More generally, a linear map $A : V \to V$ on a finite-dimensional inner product space $(V, \langle \cdot, \cdot \rangle)$ is **symmetric** or **self-adjoint** if $A^T = A$; that is, if the following holds:
$$\langle Av, w \rangle = \langle v, Aw \rangle$$
for all $v, w \in V$.

Exercise 11. Suppose A is an $n \times n$ self-adjoint matrix. Let $x \in \mathbb{C}^n$ be a nonzero vector such that $Ax = \lambda x$ for some $\lambda \in \mathbb{C}$ (we still require A to have real entries). Show that λ is real and the real part of x is an eigenvector of A. (Hint: Consider $\overline{x}^T Ax$).

Definition 2.2. An $n \times n$ matrix A is **orthogonally diagonalizable** if it is diagonalizable in the form $A = PDP^{-1}$, where P is an orthogonal matrix. A linear map $A : V \to V$ on a finite-dimensional inner product space is **orthogonally diagonalizable** if there is a matrix representing it that is orthogonally diagonalizable.

Exercise 12. Show that if an $n \times n$ matrix is orthogonally diagonalizable, then it is self-adjoint.

Exercise 13. Suppose $A = PRP^{-1}$ with P orthogonal and R upper triangular. Show that if A is symmetric, then R is diagonal.

Exercise 14. Suppose that A is an $n \times n$ matrix that is diagonalizable in the form PDP^{-1}. Show that any eigenvalue shows up in the diagonal matrix D the same number of times as its geometric multiplicity.

Definition 2.3. The collection of eigenvalues of a linear map on a finite-dimensional vector space is oftened called its **spectrum**.

The following theorem is a classic. Halmos' discussion about it in [Hal58, Sec. 79] is great, I really recommend it. (Note he states the theorem a little differently, i.e. in terms of projections.) Also, the theorem can be rephrased to be about self-adjoint linear maps on finite-dimensional inner product spaces in the ovious way.

Theorem 2.4. (**The Spectral Theorem**.) Let A be an $n \times n$ symmetric matrix. Then:

1. The spectrum of the matrix A has n real eigenvalues, counting algebraic multiplcities.

2. The eigenspaces of the matrix A are mutually orthogonal.

3. The matrix A is orthogonally diagonalizable.

4. The algebraic and geometric multiplicities of A are the same.

Exercise 15. Let A be an $n \times n$ symmetic matrix. Prove Theorem 2.4 using the following hints:

1. For the first statement, exercise 11 may be helpful.

2. Use eigenvectors from different eigenvalues for the second statement.

3. For the third statement, use the first statement. Also exercise 13 and real Schur decompositions may be useful.

4. Exercise 14 may be helpful for the last statement.

Exercise 16. Suppose A and B are orthogonally diagonalizable matrices that commute. Show that AB is orthogonally diagonalizable. Take a moment to appreciate why the Spectral Theorem makes showing this so much easier.

Definition 2.5. Let A be an $n \times n$ symmetric matrix, a **spectral decomposition** for A is an expression of the form:

$$A = \sum_{i=1}^{n} \lambda_i u_i u_i^T$$

where $\lambda_1, ..., \lambda_n$ are the eigenvalues of A and $u_1, ..., u_n$ are orthonormal eigenvectors.

Exercise 17. Let A be an $n \times n$ symmetric matrix.

1. Show that the matrix A has a spectral decomposition.
2. Given any unit vector $u \in \mathbb{R}^n$, define the matrix $B := uu^T$. Note B is a symmetric matrix. Show that this is an orthogonal projection onto some subspace. Specify the subspace.
3. Use part 2 to interpret the spectral decomposition of Definition 2.5 in terms of projections.

Exercise 18. Obtain a spectral decomposition of the matrix:

$$\begin{pmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{pmatrix}$$

Exercise 19. Let $\{u_1, ..., u_n\}$ be an orthonormal basis of $\mathbb{R}^n$ and let $\lambda_1, ..., \lambda_n$ be real scalar. Define the matrix:

$$A := \sum_{i=1}^{n} \lambda_i u_i u_i^T$$

Show that A is symmetric and that the eigenvalues of A are $\lambda_1, ..., \lambda_n$.

Definition 2.6. Let A be an $n \times n$ symmetric matrix. A **quadratic form** on $\mathbb{R}^n$ is a function of the form:

$$Q : \mathbb{R}^n \to \mathbb{R}, \qquad Q(x) := x^T A x$$

Exercise 20. Consider the following quadratic forms and write their corresponding matrices:

1. $Q : \mathbb{R}^n \to \mathbb{R}, \qquad Q(x) = ||x||^2$
2. $Q : \mathbb{R}^2 \to \mathbb{R}, \qquad Q(x, y) = 3x^2 - 4xy + 7y^2$
3. $Q : \mathbb{R}^3 \to \mathbb{R}, \qquad Q(x, y, z) = 5x^2 + 3y^2 + 2z^2 - xy + 8yz$

Theorem 2.7. (**Principal Axes Theorem**.) Given a quadratic form $Q : \mathbb{R}^n \to \mathbb{R}$, there is an orthogonal change of variables $y = Px$ that gets rid of the cross-product terms.

Definition 2.8. Given a quadratic form $Q : \mathbb{R}^n \to \mathbb{R}$. Let $\{v_1, ..., v_n\}$ be a basis of $\mathbb{R}^n$ such that Q has no cross-product terms in terms of this basis. The spans $\text{span}(v_i)$, are called the **principal axes** of the quadratic form.

Exercise 21. Prove Theorem 2.7. (Hint: What do the cross-product terms correspond to in the matrix of the form? Note the matrix of the form is symmetric.)

Exercise 22. Get rid of the cross-product term in the quadratic form $Q(x, y) = x^2 - 8xy - 5y^2$.

Definition 2.9. Let $Q : \mathbb{R}^n \to \mathbb{R}$ be a quadratic form. Then:

1. Q is **positive definite** if $Q(x) > 0$ for all $x \neq 0$.

2. Q is **positive semidefinite** if $Q(x) \geq 0$ for all x.

3. Q is **negative definite** if $Q(x) < 0$ for all $x \neq 0$.

4. Q is **negative semidefinite** if $Q(x) \leq 0$ for all x.

5. Q is **indefinite** if $Q(x)$ is none of the above.

Theorem 2.10. Let A be an $m \times n$ symmetric matrix and consider the quadratic form $Q(x) = x^T Ax$. Then:

1. Q is positive definite if and only if the spectrum of A is positive (all eigenvalues are positive).

2. Q is negative definite if and only if the spectrum of A is negative (all eigenvalues are negative).

3. Q is indefinite if and only if the spectrum of A has both positive and negative eigenvalues.

Exercise 23. Prove Theorem 2.10. (Hint: Apply Theorem 2.7.)

Exercise 24. We say a matrix A has the properties in Definition 2.9 if the quadratic form $Q(x) = x^T Ax$ has them.

1. Show that $B^T B$ is positive semidefinite, where B is an $m \times n$ matrix.

2. Show that if B is an invertible $n \times n$ matrix, then $B^T B$ is positive definite.

Exercise 25. Show that if A is an $n \times n$ positive definite symmetric matrix, then there exists a positive definite matrix B such that $A = B^T B$. (Hint: Use that A is orthogonally diagonalizable with diagonal matrix D. Write $D = C^T C$ for some matrix C and let $B = PCP^T$.)

Definition 2.11. Let A be an $n \times n$ matrix. A **Cholesky decomposition** of A is a factorization $A = R^T R$, where R is upper triangular with positive entries on the diagonal.

Exercise 26. Show that an $n \times n$ matrix A has a Cholesky decomposition if and only if it is poistive definite. (Hint: QR factorization and exercise 25).

Exercise 27. Let A be an $n \times n$ invertible symmetric matrix. Show that if A is positive definite, then so is A^{-1}.

Exercise 28. Let D be the $n \times n$ diagonal matrix with the numbers $\lambda_1 \geq \lambda_2 \geq .... \geq \lambda_n$ on its diagonal in order from greatest to lowest from left to right. Show that the quadratic form $Q : \mathbb{R}^n \to \mathbb{R}$ defined by $Q(x) = x^T Dx$ is such that $Q(x) \leq \lambda_1$ for all x in the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n \mid x^T x = 1\}$.

Exercise 29. Let A be an $n \times n$ symmetric matrix. Let Q_A be the corresponding quadratic form and let $S^{n-1} = \{x \in \mathbb{R}^n \mid x^T x = 1\}$ be the unit sphere in $\mathbb{R}^n$.

1. Apply Theorem 2.7 to make an orthogonal change of variables so that the quadratic form becomes the quadratic form Q_D given by a matrix as in exercise 28 with the entries the eigenvalues of A and the columns of the change of variable $y = Px$ are corresponding orthonormal eigenvectors. Show that Q_A and Q_D obtain the same values on the unit sphere S^{n-1}.

2. Use the previous step and exercise 28 to show that Q_A obtains the maximum λ_1 on S^{n-1}.

This shows that Q_A obtains the largest eigenvalue as its maximum when constrained to the unit sphere, and that it does so at a unit eigenvector of A. An analogous argument shows that it obtains the smallest eigenvalue as its minimum when constrained to the unit sphere.

A similar approach to the one of the previous exercise can be done to prove the following theorem:

Theorem 2.12. Let A be an $n \times n$ symmetric matrix and let Q be the corresponding quadratic form. Let $\lambda_1 \geq \lambda_2 \geq .... \geq \lambda_n$ be the eigenvalues of A and $u_1, ..., u_n$ be corresponding unit eigenvectors. Then for any integer k with $1 \leq k \leq n$, the form Q constrained to:

$$x^T x = 1 \quad x^T u_1 = 0 \quad x^T u_2 = 0 \quad ... \quad x^T u_{k-1} = 0$$

obtains the maximum λ_k at the eigenvector u_k. (Note the constraint $x^T u_j = 0$ means the hyperplane defined by u_j of vectors orthogonal to u_j.)

Exercise 30. Consider the matrix:

$$A := \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$$

Then:

1. When is the function $x \mapsto ||Ax||^2$ maximized when constrained to S^2? (Hint: Consider the quadratic form $x \mapsto x^T A^T Ax$.)

2. What is the image of the unit sphere S^2 under the linear map A?

3 The Singular Value Decomposition

Exercise 31. Let A be an $m \times n$ matrix. Show that the eigenvalues of $A^T A$ are all nonnegative. (Hint: Let $v_1, ..., v_n$ be orthonormal eigenvectors for the eigenvalues $\lambda_1, ..., \lambda_n$of the matrix $A^T A$ and consider the quadratic form corresponding to $A^T A$.)

Definition 3.1. Let A be an $m \times n$ matrix. Let $\lambda_1 \geq ... \geq \lambda_n \geq 0$ be the eigenvalues of $A^T A$. The **singular values** of A are the numbers $\sigma_i := \sqrt{\lambda_i}$, for $i = 1, ..., n$.

Exercise 32. Let A be an $m \times n$ matrix, and let $v_1, ..., v_n$ be orthonormal vectors corresponding to the eigenvalues $\lambda_1, ..., \lambda_n$ of the matrix $A^T A$. What is the geometric relationship between the vectors $||Av_i||$ and the singular values $\sigma_i = \sqrt{\lambda_i}$?

Exercise 33. Let A be an $m \times n$ matrix. Show that eigenvectors corresponding to different singular values are orthogonal.

Exercise 34. Let A be an $m \times n$ matrix and let $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n \geq 0$ be its singular values. Suppose there are exactly r nonzero singular values $\sigma_1, ..., \sigma_r$ with corresponding orthonormal eigenvectors $v_1, ..., v_r$. Show that $\{Av_1, ..., Av_r\}$ is an orthogonal basis for $\mathrm{im}(A)$ and so the rank of A is r. (Hint: Complete the basis $\{v_1, ..., v_r\}$ to an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors and check directly.)

Definition 3.2. Let A be an $m \times n$ matrix. A **singular value decomposition** is a factorization of the form:

$$A = U\Sigma V^T$$

where U is an orthogonal $m \times m$ matrix, V is an orthogonal $n \times n$ matrix, D is an $r \times r$ matrix where r is the rank of A, and Σ is of the form:

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$$

with D a diagonal matrix with positive entries on its diagonal and with the lowest right block Σ having size $(m-r) \times (n-r)$.

Theorem 3.3. Let A be an $m \times n$ matrix of rank r and let $\sigma_1 \geq .... \geq \sigma_r > 0$ be the first r singular values of A. Let Σ be a matrix of the form:

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$$

where D is of size $r \times r$ and is diagonal with entries the singular values in decreasing order. There exist an orthogonal $m \times m$ matrix U and an orthogonal $n \times n$ matrix V such that:

$$A = U\Sigma V^T$$

Exercise 35. Prove Theorem 3.3 by considering the following algorithm:

1. Let $\lambda_1 \geq ... \geq \lambda_r$ be the first nonzero eigenvalues of $A^T A$ and $v_1, ..., v_r$ be corresponding orthonormal eigenvectors. Complete this to an orthonormal eigenvector basis $\{v_1, ..., v_n\}$ of $\mathbb{R}^n$. Thus, by Exercise 34, we know $Av_1, ..., Av_r$ is an orthogonal basis of $\operatorname{im}(A)$. Normalize to obtain an orthonormal basis $\{u_1, ..., u_m\}$ of $\mathbb{R}^m$.

2. Define the matrices U and V to have as columns the basis vectors $\{u_1, ..., u_m\}$ and $\{v_1, ..., v_n\}$ respectively.

3. Let D be a diagonal matrix with the first r singular values. Let Σ be of the form given in Definition 3.2.

This gives the SVD, show it works by doing the following:

1. Show that:

$$AV = \begin{pmatrix} Av_1 & ... & Av_r & 0 & ... & 0 \end{pmatrix} = \begin{pmatrix} \sigma_1 u_1 & ... & \sigma_r u_r & 0 & ... & 0 \end{pmatrix}$$

2. Show that $U\Sigma = AV$.

3. State that U and V are orthogonal and finish the proof.

Exercise 36. Use the above algorithm, to compute an SVD for the matrices:

$$\begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} \qquad \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix}$$

Exercise 37. Show that if A is an $m \times n$ matrix with SVD decomposition $A = U\Sigma V^T$, then the columns of V are eigenvectors of A^TA and the columns of U are eigenvectors of AA^T. Show that the diagonal entries of Σ are the singular values of A. (Hint: Use the SVD for the matrices A^TA and AA^T.)

Exercise 38. Let A be an $m \times n$ matrix. Show that A is invertible if and only if A has n nonzero singular values.

Definition 3.4. Let A be an $n \times n$ matrix of rank r. A **reduced singular value decomosition** of A is a factorization $A = U_rDV_r^T$, where D be an $r \times r$ diagonal matrix with positive entries, U_r is an $m \times r$ matrix with orthogonal columns, and V_r is an $n \times r$ matrix with orthogonal columns. The **Moore-Penrose inverse** or **pseudoinverse** of A is the matrix:

$$A^+ = V_rD^{-1}U_r^T$$

Exercise 39. Obtain a reduced singular value decomposition from an SVD of A for each of the matrices in exercise 36.

Exercise 40. Let A be an $m \times n$ matrix of rank r with Moore-Penrose inverse $A^+ = V_rD^{-1}U_r^T$. Show that:

1. The linear map AA^+ is the projection of $\mathbb{R}^m$ onto $\mathrm{im}(A)$.
2. The linear map A^+A is the projection of $\mathbb{R}^n$ onto $\mathrm{im}(A^T)$
3. $AA^+A = A$ and $A^+AA^+ = A^+$
4. Let $b \in \mathbb{R}^m$ be a vector. Show that A^+b gives the least-squares solution to $Ax = b$.

References

[Hal58] P.R. Halmos. Finite-Dimensional Vector Sapces. Reprinting of the 1958 second edition. *Undergraduate Texts in Mathematics.* Springer-Verlag, New York-Heidelberg, 1974.

[L94] D.C. Lay. Linear Algebra and its Applications. Fourth Edition. Addison-Wesley, 2012.

[Rom08] S. Roman. Advanced Linear Algebra. Third Edition. *Graduate Texts in Mathematics, 135.* Springer, New York, 2008.