Lab 7

DSGA-1014: Linear Algebra and Optimization

CDS at NYU
Zahra Kadkhodaie

Fall 2021

Diagonalization

Let A be a square $n \times n$ matrix with n linearly independent eigenvectors. Then A is called diagonalizable and we can write

$$A = Q \Lambda Q^{-1}$$

where Q consists of eigenvectors of A and Λ is a diagonal matrix consists of eigenvalues of A. Note that Q^{-1} exits, because eigenvectors are linearly independent, so Q is invertible.

$$Ax = \underbrace{Q \underbrace{\Lambda \underbrace{Q^{-1} x}_{\text{change of basis: represent } x \text{ in eigenspace coordinates}}}_{\text{multiply by a diagonal matrix: rescale}}}_{\text{change of basis: go back to canonical basis}}$$

Diagonalization

Not all square matrices are diagonalizable:

1. If A has n distinct eigenvalues, then all the n eigenvectors are independent and A is diagonalizable. That is, $\dim(E_{\lambda_i}(A)) = 1$ for $\forall i$
2. If A has repeated eigenvalues, then it *might* be the case that A does not have n linearly independent eigenvectors. That is for some λ_i, $\dim(E_{\lambda_i}(A)) <$ number of times λ_i is repeated. In this case, A is not diagonalizable, because Q does not have an inverse.
3. If A is symmetric, then it is guaranteed that A has n linearly independent eigenvectors and is diagonalizable (even if some eigenvalues are repeated!). Moreover, not only eigenvectors are independent, but they are orthogonal too. So

$$A = Q\Lambda Q^T$$

   This is referred to as Spectral Theorem.

1. Let $P \in R^{n \times n}$ be a projection matrix.

(a) Show that P is always diagonalizable.

(b) What are the eigen values?

(c) Is P orthogonal? NO - Counter example

(d) Define $P = \begin{bmatrix} .2 & .4 & 0 \\ .4 & .8 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. For each eigen value, give the dimension of corresponding eigen space, $E_\lambda(P)$.

a) $\underset{n \times n}{P} = VV^T$ Columns of V: k orthogonal vectors $V \in R^{n \times k}$

$P^T = (VV^T)^T = VV^T = P \Rightarrow P$ is symmetric $\Rightarrow$ diagonalizable.

b) V is an orthogonal basis for projection subspace.

we define U as an orthogonal basis for the complement.

Concatinate V and U to get an orthogonal basis for R^n.

Define Q as concatination of v and u. Define Λ as a diagonal matrix, with $\lambda_i = 1$ for $1 \le i \le k$ and $\lambda_i = 0$ for $k < i \le n$:

$$P = vv^T = Q \Lambda Q^T = \Big[\ \underbrace{|\,|\,|\,|}_{v}\ \underbrace{|\,|}_{u}\ \Big] \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & 0 \end{bmatrix} \begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{bmatrix} \begin{array}{l} \}\, v \\ \\ \}\, u \end{array}$$

d) $\operatorname{rank}(P) = 2 \implies \lambda_1 = \lambda_2 = 1$ and $\lambda_3 = 0$

$\dim(E_{\lambda=1}(P)) = 2$ $\qquad$ $\dim(E_{\lambda=0}(P)) = 1$

Alternatively use: $\operatorname{Tr}(P) = \sum \lambda_i$

$$.2 + .8 + 1 = 2 = \lambda_1 + \lambda_2 + \lambda_3$$

from (b) we know that λ is either 1 or zero. So $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = 0$

2. What matrix A has eigenvalues $\lambda = 1, -1$ and eigenvectors $v_1 = (\cos\theta, \sin\theta)$ and $v_2 = (-\sin\theta, \cos\theta)$? Which of these properties hold? $A = A^T$, $A^2 = I$, $A^{-1} = A$.

$A \in \mathbb{R}^{2\times 2}$

$\langle v_1, v_2 \rangle = 0$ and $\|v_1\| = \|v_2\| = 1 \Rightarrow \{v_1, v_2\}$ orthonormal set.

$\Rightarrow A = Q \Lambda Q^{-1} = Q \Lambda Q^T = (Q \Lambda Q^T)^T = A^T$

$\Rightarrow A = A^T$ symmetric

$A^2 = Q \Lambda Q^T Q \Lambda Q^T = Q \Lambda^2 Q^T = Q I Q^T$ because $\lambda = 1, -1$

$\Rightarrow A^2 = I$

$A^{-1} = (Q \Lambda Q^T)^{-1} = Q \Lambda^{-1} Q^T = Q \Lambda Q^T$ because $\lambda = 1, -1$

$\Rightarrow A^{-1} = A$

$A^{-1} = A^T \Rightarrow A$ is orthogonal. Also we showed it is symmetric.

Symmetric orthogonal matrices are very restricted. Their eigenvalues are $\lambda_i = \{-1, 1\}$.

If $\lambda_1 = \lambda_2 = 1 \Rightarrow A = \begin{bmatrix} \cos & -\sin \\ \sin & \cos \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos & \sin \\ -\sin & \cos \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$

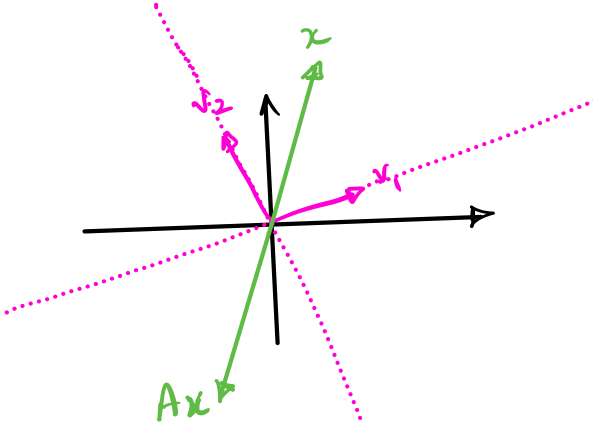
or a rotation with $\theta = 0$

If $\lambda_1 = \lambda_2 = -1 \Rightarrow A = \begin{bmatrix} \cos & -\sin \\ \sin & \cos \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos & \sin \\ -\sin & \cos \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

A is a rotation with $\theta = \pi$

If $\lambda_1 = 1, \lambda_2 = -1 \Rightarrow A = \begin{bmatrix} \cos & -\sin \\ \sin & \cos \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos & \sin \\ -\sin & \cos \end{bmatrix} = \overbrace{\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}}^{\text{symmetric \& orthogonal}}$

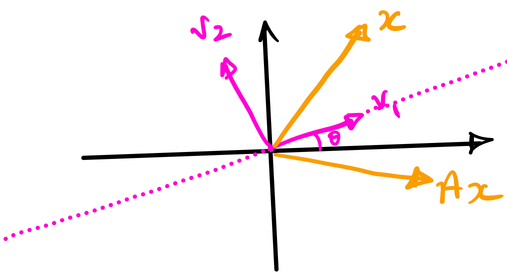
A is reflection around first eigen vector



rotation $\theta = \pi$

vectors in span(v_1) are reversed.

vectors in span(v_2) are reversed.



reflection around first eigenvector:

vectors in span(v_1) are unchanged.

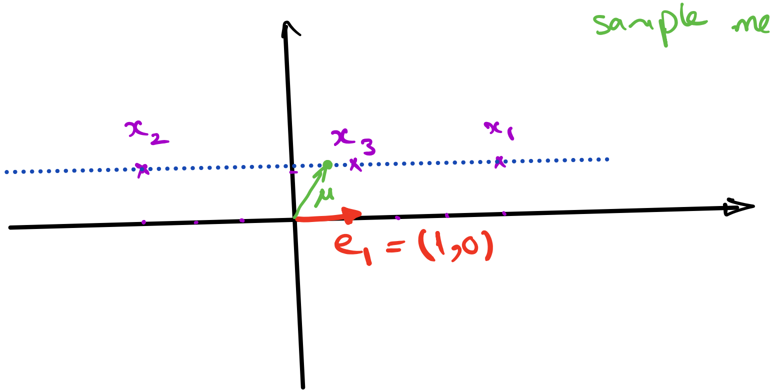
vectors in span(v_2) are reversed.

3. Define $x_1 = (4,1), x_2 = (-3,1)$, and $x_3 = (1,1)$

(a) Give a one-dimensional affine subspace of R^2 that best approximates these three points.

(b) Use this to represent each point using a single number (i.e., reduce the dimension from 2 to 1).

(c) Describe the eigen decomposition of the covariance matrix without computing it directly.



sample mean: $\vec{\mu} = (\frac{2}{3}, 1)$

affine subspace: $\text{span}(\vec{e}_1) + \vec{\mu}$

b) $x_1 = (4 - 2/3)\vec{e}_1 + \vec{\mu}$

$x_2 = (-3 - 2/3)\vec{e}_1 + \vec{\mu}$

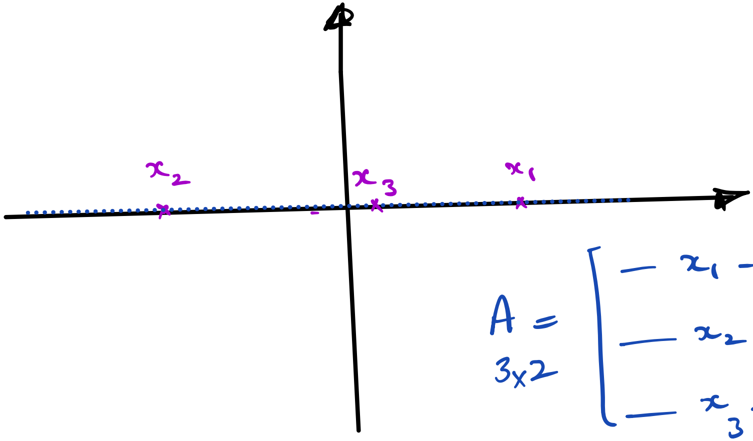
$x_3 = (1 - 2/3)\vec{e}_1 + \vec{\mu}$

$\Rightarrow$

$\hat{x}_1 = \frac{10}{3}$

$\hat{x}_2 = \frac{-11}{3}$

$\hat{x}_3 = 1/3$



mean-centered data

$$\underset{3\times2}{A} = \begin{bmatrix} — & x_1-\mu & — \\ — & x_2-\mu & — \\ — & x_3-\mu & — \end{bmatrix} \Rightarrow \underset{2\times2}{S} = A^T A \quad \text{: covariance matrix}$$

After removing mean, projection of these samples onto span of $\vec{e}_1$ is equal to the original samples (lossless). That is variance in the direction of $\vec{e}_2 = 0$

$S = Q D Q^T$

$\lambda_1 > 0$ variance along e_1

$\lambda_2 = 0$

first eigenvector $v_1 = \vec{e}_1$

$v_2 = e_2$

4. Suppose $A \in R^{n \times n}$ has a linearly independent list of n eigenvectors $v_1, \ldots, v_n$ with real eigenvalues $\lambda_1, \ldots, \lambda_n$. Can we factor A in a way similar to the spectral decomposition? Show that if $v_1, \ldots, v_n$ are orthonormal, then A has to be symmetric.

we can write $A = V \Lambda V^{-1}$

V is not neccessarily orthogonal here.

If V is orthogonal : $V^{-1} = V^T \Rightarrow$

$A = V \Lambda V^T$

$A^T = (V \Lambda V^T)^T = V \lambda V^T = A$

symmetric A.

5. Let A and B be diagonalizable matrices. Also assume that α is an eigenvalue of A and β is an eigenvalue of B. Under what condition $\alpha\beta$ is an eigenvalue of AB?

If A and B share the eigen vector associated with α and β:

$$ABv = A\beta v = \beta A v = \beta\alpha v$$

Note: Generally if A and B commute then they share eigenvector.