Recitation 2

Alex Dong

CDS, NYU

Fall 2020

Concept Review: Orthogonal Matrices

- **Orthogonal** matrices have *orthonormal* columns
  - Stronger condition than having orthogonal columns
  - Bad terminology thats grandfathered in
- We will see a lot of these matrices
- Orthogonal matrices preserve angles and norms
  - This leads to a very natural *change of basis* - more later

Questions: Orthogonal Matrices

1. Let $Q \in \mathbb{R}^{n \times n}$ be an orthogonal matrix. Let $x, y \in \mathbb{R}^n$.
   i. Show that $\|Qx\| = \|x\|$.
   ii. Show that $\langle Qx, Qy \rangle = \langle x, y \rangle$.

Solutions 1: Orthogonal Matrices

1. Let $Q \in \mathbb{R}^{n \times n}$ be an orthogonal matrix. Let $x, y \in \mathbb{R}^n$.
   i. Show that $\|Qx\| = \|x\|$.

Solution

Note: Recall the "lin. comb. of columns" method of matrix multiplication.

Let $Q = \begin{bmatrix} | & & | \\ \mathbf{q_1} & \dots & \mathbf{q_n} \\ | & & | \end{bmatrix}$ *and* $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. *Then* $Qx = \begin{bmatrix} | \\ \sum_{i=1}^n x_1 \mathbf{q_i} \\ | \end{bmatrix}$, *and*

$$\|Qx\| = \langle \textstyle\sum_{i=1}^n x_i \mathbf{q_i}, \sum_{i=1}^n x_i \mathbf{q_i} \rangle$$
$$\|Qx\| = \textstyle\sum_{i=1}^n \langle\ x_i \mathbf{q_i}, x_i \mathbf{q_i} \rangle + 2 \sum_{i \neq j} \langle\ x_i \mathbf{q_i}, x_j \mathbf{q_j} \rangle$$
$$\|Qx\| = \textstyle\sum_{i=1}^n x_i^2 \|q_i\| \qquad \textit{by orthogonality of } q_i$$
$$\|Qx\| = \textstyle\sum_{i=1}^n x_i^2 \qquad \textit{by normality of } q_i.$$
$$\|Qx\| = \|x\|.$$

Solutions 2: Orthogonal Matrices

1. Let $Q \in \mathbb{R}^{n \times n}$ be an orthogonal matrix. Let $x, y \in \mathbb{R}^n$.
    ii. Show that $\langle Qx, Qy \rangle = \langle x, y \rangle$.

Solution

$\langle Qx, Qy \rangle = x^T Q^T Q y = x^T I y = x^T y = \langle x, y \rangle$

Concept Review: Gram-Schmidt Process

- Gram-Schmidt Process turns a basis of linearly independent vectors into orthonormal vectors
- Understanding the GS process is important, but we will mainly only use its existence
  - Let $v_1, ..., v_n$ be a basis ...
    ... and by GS process, let $u_1, ..., u_n$ be orthonormal with $Span(v_1, ..., v_n) = Span(u_1, ..., u_n)$:
  - Let $u_1, ..., u_n$ be an orthonormal basis of $\mathbb{R}^n$
- Related to QR Factorization

Questions: GS Process and QR Factorization

1. Let $A \in \mathbb{R}^{n \times n}$ have linearly independent columns. Show that there is a matrix $Q \in \mathbb{R}^{n \times m}$ and $R \in \mathbb{R}^{n \times n}$ s.t that $A = QR$, where Q has orthonormal columns and R is upper triangular.

   (Hint: Recall the "linear combination of columns interpretation of matrix multiplication").

1. Let $A \in \mathbb{R}^{m \times n}$ have linearly independent columns. Show that there is a matrix $Q \in \mathbb{R}^{m \times m}$ and $R \in \mathbb{R}^{m \times n}$ s.t that $A = QR$, where Q has orthonormal columns and R is upper triangular.

Solution

First, let $v_1, ..., v_n$ be the columns of A.
Apply the GS process to get $u_1, ..., u_n$.
Now, let Q have $u_1, ..., u_n$ as its columns.
Note that by the GS process, we have
$Span(v_1, ..., v_i) = Span(u_1, ..., u_i)\ \forall i \in \{1, ..., n\}$.
Then each column v_i is a linear combination of the columns $u_1, ..., u_i$.
Then this exactly saying that $A = QR$, where R contains the coefficients that transforms $u_1, ..., u_i$ into $v_1, ..., v_i$ $\forall i \in \{1, ..., n\}$!

More M.M.M: Linear Combination of Columns

Each column of the AB is a linear combination of the columns of A.

$$\begin{bmatrix} | & | & \dots & | & | \\ \mathbf{a_1} & \mathbf{a_2} & \dots & \mathbf{a_{k-1}} & \mathbf{a_k} \\ | & | & \dots & | & | \end{bmatrix} \begin{bmatrix} b_{1,1} & \dots & b_{1,m} \\ b_{2,1} & \dots & b_{2,m} \\ \vdots & \vdots & \vdots \\ b_{k-1,1} & \dots & b_{k-1,m} \\ b_{k,1} & \dots & b_{k,m} \end{bmatrix}$$

$$= \begin{bmatrix} | & \dots & | \\ \sum_{i=1}^{k} \mathbf{a_i} b_{i,1} & \dots & \sum_{i=1}^{k} \mathbf{a_i} b_{i,m} \\ | & \dots & | \end{bmatrix}$$

A Note About Determinants

- Eigenvalues of a matrix can be determined by using determinants
- **Not covered in this course!**
  - "too long to define, a bit complex, and slightly useless in data science..." - Léo
- Determinants lead to a lot of cool things
  - Trace(A) = sum of eigenvalues of A (with multiplicity)
  - (&) A matrix satisfies it's own *characteristic polynomial* - Cayley Hamilton Theorem
  - (&) Matrix polynomial rabbit hole runs deep (Jordan Normal Form)
- Interesting from a pure math perspective

[0] (&) denotes extra material not covered in this course

Etymology

- *eigen*values and *eigen*vectors
- What does *eigen* mean anyway?
- German word for...
  1. own
  2. innate
  3. peculiar
  4. **intrinsic**
- A square matrix 'owns' certain vectors... or there are certain vectors that are intrinsic to a matrix.

Importance of Eigenvalues and Eigenvectors

!!! ***SERIOUSLY IMPORTANT*** !!!

- Eigen-val/vec will show up *continuously* throughout this course
- Connections to...
  - Projections and Orthogonal Projections (Lec 4)
  - Markov Chains (Lec 6)
  - Spectral Theorem (HW 6, Lec 7)
  - SVD (Lec 7)
  - Spectral Clustering (!!??) (Lec 8)
  - Positive definite and positive semi-definite matrices (Lec 10,11)
- Many other applications not covered in this course

$Av = \lambda v$. So what's the big deal?

- Square matrices are important enough to get their own name - *operators.*
- Sometimes a matrix A 'prefers' certain directions
- (!!!) These directions act as *anchors* for understanding the action of a matrix.
- We will see how to exploit these directions in order to simplify our understanding of matrices. (Lec 7)

Questions 1: Eigen

Let $A \in \mathbb{R}^{n \times n}$ have eigenvalue λ associated to eigenvector v. Show that:

1. $\forall \alpha \in \mathbb{R}$, $\lambda + \alpha$ is an eigenvalue of $A + \alpha I$ w/ eigenvector v.
2. $\forall k \in \mathbb{N}$, λ^k is an eigenvalue of A^k w/ eigenvector v.
3. Let $A \in \mathbb{R}^{n \times n}$ have eigenvalue-vector pairs $\lambda_1, ..., \lambda_n$ and $v_1, ..., v_n$.
   Also, assume that $\lambda_1 > ... > \lambda_n$.
   Prove that $v_1, ..., v_n$ are linearly independent.

   Hint: First assume that all λ_i are positive.

Solutions 1: Eigen

Let $A \in \mathbb{R}^{n\times n}$ have eigenvalue λ associated to eigenvector v. Show that:

1. $\forall \alpha \in \mathbb{R}$, $\lambda + \alpha$ is an eigenvalue of $A + \alpha I$ w/ eigenvector v.

Solution

Let $\alpha \in \mathbb{R}$, and v be an eigenvector of A.
Consider the matrix $A + \alpha I$.

$$\begin{aligned}(A + \alpha I)v &= Av + \alpha I v \\ &= \lambda v + \alpha v \\ &= (\lambda + \alpha)v\end{aligned}$$

So $\lambda + \alpha$ is an eigenvalue of $A + \alpha I$ with eigenvector v.

Solutions 1: Eigen

Let $A \in \mathbb{R}^{n \times n}$ have eigenvalue λ associated to eigenvector v. Show that:

2. $\forall k \in \mathbb{N}$, λ^k is an eigenvalue of A^k w/ eigenvector v.

Solution

Let $k \in \mathbb{N}$, and v be an eigenvector of A.
Consider the matrix A^k.

$$A^k v = A \ldots A v \qquad \textit{k times}$$

$$A^k v = A \ldots A(\lambda v) \qquad \textit{k-1 times}$$

$$A^k v = \lambda^k v$$

So λ^k is an eigenvalue of A^k with eigenvector v.

Solutions 1: Eigen

3. Let $A \in \mathbb{R}^{n \times n}$ have eigenvalue-vector pairs $\lambda_1, ..., \lambda_n$ and $v_1, ..., v_n$.
   Also, assume that $\lambda_1 > ... > \lambda_n$.
   Prove that $v_1, ..., v_n$ are linearly independent.

Solution

Let $B = A + |\lambda_n| I$. *(This is so all eigenvalues of* B *are* ≥ 0.*)*
Let $\gamma_i = \lambda_i + |\lambda_n|$ *(Problem 1, eigenvecs of* B *are also eigenvecs of* A.
Let $\alpha_1, ..., \alpha_n \in \mathbb{R}$ *s.t* $\sum_{i=1}^n \alpha_i v_i = 0$. *We will show that all* $\alpha_i = 0$.
Consider $0 = B^k(\sum_{i=1}^n \alpha_i v_i)$.

$$0 = B^k(\sum_{i=1}^n \alpha_i v_i)$$
$$0 = \sum_{i=1}^n B^k \alpha_i v_i$$
$$0 = \sum_{i=1}^n \gamma_i^k \alpha_i v_i$$
$$0 = \gamma_1^k \sum_{i=1}^n (\frac{\gamma_i}{\gamma_1})^k \alpha_i v_i$$
$$0 = \lim_{k \to \infty} \gamma_1^k \sum_{i=1}^n (\frac{\gamma_i}{\gamma_1})^k \alpha_i v_i$$
$$0 = \alpha_1 v_1 \quad \textit{since } \frac{\gamma_i}{\gamma_1} < 1 \textit{ for } i \neq 1$$

Then $0 = (\sum_{\mathbf{i=2}}^n \alpha_i v_i)$. *Repeat the previous logic to find that each* $\alpha_i v_i = 0$.
Then all $\alpha_1 = 0$. *So* $v_1, ..., v_n$ *are linearly independent.*

Questions 2: Properties of Orthogonal Matrices

Let $Q \in \mathbb{R}^{n \times n}$ be orthogonal.

1. Does Q necessarily have eigenvalues and eigenvectors?

Assume that Q has eigenvalues $\lambda_1, ..., \lambda_k$.

2. Describe the eigenvalues of Q.

Solutions 2: Properties of Orthogonal Matrices

Let $Q \in \mathbb{R}^{n \times n}$ be orthogonal.

1. Does Q necessarily have eigenvalues and eigenvectors?

Solution

No, consider the matrix $Q = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ *(90 deg CCW rotation in* $\mathbb{R}^2$*).*

Assume that Q has eigenvalues $\lambda_1, ..., \lambda_k$.

2. Describe the eigenvalues of Q.

Solution

Since Q *is orthogonal then* $\forall x \in \mathbb{R}^n$

$\|Qx\| = \langle Qx, Qx \rangle$

$\|Qx\| = x^T Q^T Qx$

$\|Qx\| = xIx$

$\|Qx\| = \|x\|$

Now, if x *is an eigenvector of* Q *with eigenvalue* λ*, then we have*

$\|x\| = \|Qx\| = \|\lambda x\| = |\lambda| \|x\|$. *So* $\lambda = \pm 1$.