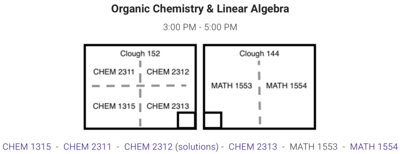
Announcements

Tuesday, April 24

- Please fill out the CIOS form online. *Current response: 68%*
  - If we *get an 80% response rate* before the final, I'll drop the *two* lowest quiz grades instead of one.

- **Resources**
  - Office hours: posted on the website.
  - Math Lab at Clough is also a good place to visit.
  - Materials to review: `https://people.math.gatech.edu/~leslava3/1718S-2802.html`
  - Reading day Wednesday, April 25th:



- **Final Exam:**
  - Date: Thursday, April 26th
  - Location: This lecture room, College of Comp. 017
  - Time: 2:50-5:40 pm

First factorization of MATH 2802

A guru provides, for (suitable) $m \times n$ matrix A, matrices L and U such that

- L is lower triangular $m \times m$ matrix (with *ones on the diagonal*)
- U is an $m \times n$ row echelon form (not necessary reduced)
- $A = LU$

E.g.

$$A = \begin{pmatrix} 2 & 4 & -1 & 5 & -1 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix}$$

The matrix LU

It helps to interpret L as *instructions on how to sum rows* of U.

A short cut to solving equations

How to save time in solving $Ax = b$?

1. Visit the guru and *get L and U*,
2. *Quickily* solve $Ly = b$,
3. *Quickily* solve $Ux = y$,
4. Claim that x **is a solution** to $Ax = b$.

Are we **allowed to** do that?

$A = LU$ Is x really a solution?

Plugin values $Ax = L\underbrace{Ux}_{y} = \underbrace{Ly}_{b} = b$

Construct the matrix L

~~$L = ((E_6)(E_5E_4)(E_3E_2E_1))^{-1}$~~

2. Separate the row reduction according to 'clearing' pivot columns

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} = A_1$$

$$\sim A_2 = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} = U$$

$$\begin{bmatrix} 2 \\ -4 \\ 2 \\ -6 \end{bmatrix} \begin{bmatrix} 3 \\ -9 \\ 12 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} \begin{bmatrix} 5 \end{bmatrix}$$

$\div 2$ ↓ $\div 3$ ↓ $\div 2$ ↓ $\div 5$ ↓

$$\begin{bmatrix} 1 & & & \\ -2 & 1 & & \\ 1 & -3 & 1 & \\ -3 & 4 & 2 & 1 \end{bmatrix}, \quad \text{and} \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{bmatrix}$$

A second example

Find the LU factorization of A:

$$A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix} \sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & -3 & -1 & 6 \\ 0 & 6 & 2 & -7 \\ 0 & -9 & -3 & 13 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 10 \end{bmatrix} \sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U$$

$$\begin{bmatrix} 2 \\ 6 \\ 2 \\ 4 \\ -6 \end{bmatrix} \quad \begin{bmatrix} 3 \\ -3 \\ 6 \\ -9 \end{bmatrix} \quad \begin{bmatrix} 5 \\ -5 \\ 10 \end{bmatrix}$$

$$\div 2 \qquad \div 3 \qquad \div 5$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

$$\begin{bmatrix} 1 & & & \\ 3 & 1 & & \\ 1 & -1 & 1 & \cdots \\ 2 & 2 & -1 & \\ -3 & -3 & 2 & \end{bmatrix},$$

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 \\ 2 & 2 & -1 & 1 & 0 \\ -3 & -3 & 2 & 0 & 1 \end{bmatrix}$$

Section 5.1

Eigenvectors and Eigenvalues

Eigenvectors and Eigenvalues

Definitions

If *v* **is not zero** and $Av = \lambda v$ then
v is *an eigenvector* and λ is *its eigenvalue*.

- Eigenvalues and eigenvectors are only for square matrices.
- Eigenvectors are by definition nonzero. *Eigenvalues may be equal to zero.*

Example

$$A = \begin{pmatrix} 2 & 2 \\ -4 & 8 \end{pmatrix} \qquad v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Multiply:

$$Av = \begin{pmatrix} 2 & 2 \\ -4 & 8 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = 4v$$

Hence v is an *eigenvector* of A, with *eigenvalue* $\lambda = 4$.

Eigenspaces

The **λ-eigenspace** is a *subspace* of $\mathbf{R}^n$ containing all *eigenvectors of A with eigenvalue λ, plus the zero* vector:

$$\lambda\text{-eigenspace} = \mathrm{Nul}\big(A - \lambda I\big).$$

Find a basis for the 2-eigenspace of

λ

$$A = \begin{pmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{pmatrix}.$$

$$A - 2I = \begin{pmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{pmatrix} \xrightarrow{\text{row reduce}} \begin{pmatrix} 1 & -\frac{1}{2} & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{\text{parametric vector form}} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = v_2 \begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \end{pmatrix} + v_3 \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$$

$$\xrightarrow{\text{basis}} \left\{ \begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Summary

Let A be an $n \times n$ matrix and let λ be a number.

1. λ is an **eigenvalue of A**
   if and only if $(A - \lambda I)x = 0$ *has a nontrivial solution*.

2. Finding a basis for the **λ-eigenspace of A**
   means finding a basis for $\mathrm{Nul}(A - \lambda I)$,

3. The **eigenvectors** with eigenvalue λ are
   the nonzero elements of $\mathrm{Nul}(A - \lambda I)$

- If we *know λ is eigenvalue*: easy to find eigenvectors (row reduction).
- And to **find all eigenvalues**? Will need to *compute a determinant*.
  Finding λ that has a non-trivial solution to $(A - \lambda I)v = 0$ boils down to finding λ that makes $\det(A - \lambda I) = 0$.

Some facts you can work out yourself

Fact 1

A is **invertible** if and only if 0 *is not an eigenvalue* of A.

Fact 2

If $v_1, v_2, \ldots, v_k$ are eigenvectors of A with **distinct eigenvalues** $\lambda_1, \ldots, \lambda_k$, then $\{v_1, v_2, \ldots, v_k\}$ is *linearly independent*.

Consequence of Fact 2

An $n \times n$ matrix has **at most n** distinct eigenvalues.

Fact 3

The **eigenvalues** of a triangular matrix are the *diagonal entries*.

Diagonalization

The Diagonalization Theorem

An $n \times n$ matrix A is **diagonalizable** if and only if A has n *linearly independent eigenvectors*.

In this case, $A = PDP^{-1}$ for

$$P = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \qquad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where $v_1, v_2, \ldots, v_n$ are linearly independent eigenvectors, and $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the *corresponding eigenvalues* (in the same order).

The Characteristic Polynomial

Last section we learn that for a square matrix A:

$$\lambda \text{ is an eigenvalue of } A \iff \det(A - \lambda I) = 0.$$

Compute Eigenvalues

The *eigenvalues* of A are **the roots** of $\det(A - \lambda I)$, which is the characteristic polynomial of A.

Definition

Let A be a square matrix. The **characteristic polynomial** of A is

$$f(\lambda) = \det(A - \lambda I).$$

The **characteristic equation** of A is the equation

$$f(\lambda) = \det(A - \lambda I) = 0.$$

How to **diagonalize a matrix** A:

1. *Find the eigenvalues* of A using the characteristic polynomial.
2. *Compute a basis* $\mathcal{B}_\lambda$ for each λ-eigenspace of A.
3. If there are fewer than n total vectors in the union of all of the eigenspace bases $\mathcal{B}_\lambda$, then the matrix is not diagonalizable.
4. *Otherwise*, the n vectors $v_1, v_2, \ldots, v_n$ in your eigenspace bases are linearly independent, and $A = PDP^{-1}$ for

$$P = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where λ_i is the eigenvalue for v_i.

Problem: Diagonalize $A = \begin{pmatrix} 4 & -3 & 0 \\ 2 & -1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$.

The *characteristic polynomial* is

$$f(\lambda) = \det(A - \lambda I) = -\lambda^3 + 4\lambda^2 - 5\lambda + 2 = -(\lambda - 1)^2(\lambda - 2).$$

Therefore the *eigenvalues are* 1 *and* 2, with respective multiplicities 2 and 1.

First compute the 1-eigenspace:

$$(A - I)x = 0 \iff \begin{pmatrix} 3 & -3 & 0 \\ 2 & -2 & 0 \\ 1 & -1 & 0 \end{pmatrix} x = 0 \xrightarrow{\text{rref}} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x = 0$$

The parametric vector form is $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = y \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Hence a basis for the 1-eigenspace is

$$\mathcal{B}_1 = \{v_1, v_2\} \quad \text{where} \quad v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Diagonalization

Example, continued

Now let's compute the 2-eigenspace:

$$(A-2I)x=0 \iff \begin{pmatrix} 2 & -3 & 0 \\ 2 & -3 & 0 \\ 1 & -1 & -1 \end{pmatrix} x = 0 \xrightarrow{\text{rref}} \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} x = 0$$

The parametric form is $x = 3z, y = 2z$, so an eigenvector with eigenvalue 2 is

$$v_3 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

Note that v_1, v_2 form a basis for the 1-eigenspace, and v_3 has a distinct eigenvalue. Thus, the eigenvectors v_1, v_2, v_3 are linearly independent and the Diagonalization Theorem says

$$A = PDP^{-1} \quad \text{for} \quad P = \begin{pmatrix} 1 & 0 & 3 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

In this case: there are 3 linearly independent eigenvectors and only 2 distinct eigenvalues.

Application

Stochastic Matrices and PageRank

Stochastic Matrices

These arise very commonly in modeling of probabalistic phenomena (Markov chains), where they are also called **transition matrices**.

Some examples:

- Matrices from the population dynamics
- Matrices from the equilibrium-prices economies

Definition

A square matrix A is **stochastic** if all of its entries are nonnegative, and the sum of the entries of each column is 1.

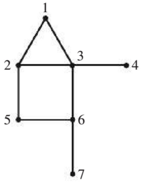
We say A is **regular** if, for some k, all entries of A^k are positive.

Definition

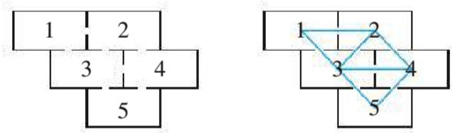
A **steady-state vector** v of A is a non-zero vector with *entries summing to 1* and such that $Av = v$.

Random walks on graphs (a.k.a Mouse on a maze)

A mouse moves freely between rooms/states = selects any with equal probability.



$$P = \begin{bmatrix} 0 & 1/3 & 1/4 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/4 & 0 & 1/2 & 0 & 0 \\ 1/2 & 1/3 & 0 & 1 & 0 & 1/3 & 0 \\ 0 & 0 & 1/4 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 0 & 0 & 1/3 & 0 \\ 0 & 0 & 1/4 & 0 & 1/2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1/3 & 0 \end{bmatrix}$$



$$P = \begin{bmatrix} 0 & 1/3 & 1/4 & 0 & 0 \\ 1/2 & 0 & 1/4 & 1/3 & 0 \\ 1/2 & 1/3 & 0 & 1/3 & 1/2 \\ 0 & 1/3 & 1/4 & 0 & 1/2 \\ 0 & 0 & 1/4 & 1/3 & 0 \end{bmatrix}$$

- *Initial state:* The mouse is located at some room i: probabilities

  $v_0 = (x_1, \vdots, x_5)$.
- Probability mouse starts at room 1 is x_1 item *Transition matrix:* $v_{n+1} = Av_n$ means that A dictates how probabilities change.
- Probability mouse is at room 3 after n steps of the walk: third entry of v_n.

Non-regular transition matrix

Disconnected states

Consider the following 'transition graph':

The transition matrix is $\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$.

Both $\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, are eigenvectors with eigenvalue 1.

So there is more than one steady-state vector!

Stochastic Matrices and Difference Equations

Through an example

Red Box has kiosks all over where you can rent movies. You can return them to any other kiosk.

- *ij entry of A:* probability that a movie rented from location j is returned to location i.

For example, if there are three locations, maybe

$$A = \begin{pmatrix} .3 & .4 & .5 \\ .3 & .4 & .3 \\ .4 & .2 & .2 \end{pmatrix}.$$

30% probability a movie rented from location 3 gets returned to location 2

On day n: x_n, y_n, z_n are the numbers of movies in locations $1, 2, 3$, respectively, and $v_n = (x_n, y_n, z_n)$.

Probabilistic Intuition

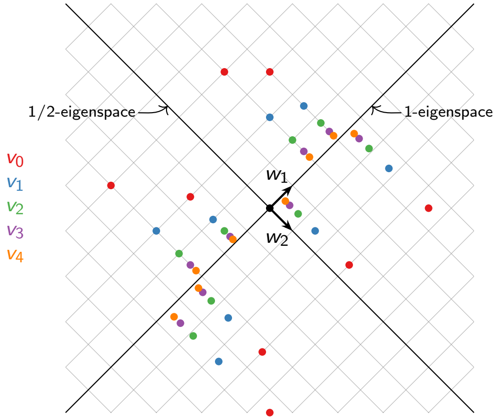
If at opening day the movies are distributed according to v_0 then, **on average**:

$$v_n = Av_{n-1} = A^2 v_{n-2} = \cdots = A^n v_0.$$

Diagonalizable Stochastic Matrices

Example, continued

Recall: $A^n = PD^nP^{-1}$ acts on the usual coordinates of v_0 in the same way that D^n acts on the $\mathcal{B}$-coordinates, where $\mathcal{B} = \{w_1, w_2\}$.



All vectors get "sucked into the 1-eigenspace."

If A is the Red Box matrix, and v_n is the vector representing the number of movies in the three locations on day n, then

$$v_{n+1} = Av_n.$$

For any starting distribution v_0 of videos in red boxes, after enough days, the distribution v ($= v_n$ for n large) is an eigenvector with eigenvalue 1:

$$Av = v.$$

In other words, eventually the number of movies in each kiosk doesn't change much.

Moreover, we know exactly what v is: a multiple of w_1

- The entries in v have to sum up to the number of intial movies (same sum as entries in v_0.

(Remember the total number of videos never changes.) Presumably, Red Box

really does have to do this kind of analysis to determine how many videos to put in each box.

Find the actual Steady State w_1

Red Box example

If one computes $\text{Nul}(A - I)$ and find that $w' = \begin{pmatrix} 7 \\ 6 \\ 5 \end{pmatrix}$

is an eigenvector with eigenvalue 1.

Then, to get a steady state, divide by $18 = 7 + 6 + 5$ to get

$$w = \frac{1}{18} \begin{pmatrix} 7 \\ 6 \\ 5 \end{pmatrix} \sim (0.39, 0.33, 0.28).$$

The long-run

So if you start with 100 total movies, eventually you'll have $100w = (39, 33, 28)$ movies in the three locations, every day.

The time spent on a state

Regardless of the intital location of a particular movie. Eventually, that movie will get 'returned' 39% of the times at location 1, 33% at location 2, and 28% at location 3.

Section 7.3

Constrained Optimization

Motivation: How to allocate resources

Problem: The government wants to repair

- w_1 hundred miles of public roads
- w_2 hundred acres of parks

Resources are limited, so cannot work on more than

- 3 miles of roads or
- 2 acres of park;
- *general condition* is:

$$4w_1^2 + 9w_2^2 \leq 36$$

How to allocate resources?
Utility function: Considering overall benefits, want to maximize

$$q(w_1, w_2) = w_1 w_2.$$

(i.e.Do not focus solely on roads nor parks)

How would you *maximize utility $q(w_1, w_2)$*?

The constraint in these optimization problems

We will keep the restriction that vectors x in $\mathbf{R}^n$ have unit length;

$$||x|| = 1, \qquad x \cdot x = 1 \qquad x^T x = 1$$

or more commonly used: $x_1^2 + x_2^2 + \cdots x_n^2$.

Example

$Q(x) = 3x_1^2 + 7x_2^2$

Plot this function in 3-dimension as:

$$\begin{pmatrix} x_1 \\ x_2 \\ Q(x) \end{pmatrix}$$

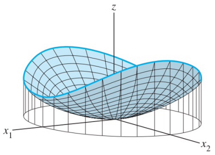


FIGURE 2 The intersection of $z = 3x_1^2 + 7x_2^2$ and the cylinder $x_1^2 + x_2^2 = 1$.

The constrained optimization problem

Given a **quadratic form** $Q(x)$, restricted to *unit vectors*,

What is the maximum and minimum values of $Q(x)$, which vectors attain such extremes?

The Constrained Optimization theorem

Theorem

Let A be a symmetric matrix and $Q(x) = x^T Ax$ a quadratic function

- **Maximum:** the maximum value of $Q(x)$ subject to $x^T x = 1$ equals the **largest eigenvalue** M of A.

  This maximum is attained by an eigenvector of A corresponding to M.
- *Minimum:* the minimum value of $Q(x)$ subject to $x^T x = 1$ equals the *smallest eigenvalue* m of A.

  This minimum is attained by an eigenvector of A corresponding to m.

How to use this information? To find maximum/minimum values of $Q(x)$, under *restriction* $x^T x = 1$:

- Find the eigenvalues of A, list them in decreasing order $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n$.
- Then maximum is $M = \lambda_1$ and minimum is $m = \lambda_n$.

Example

What is the maximum value of $Q(x) = x^T Ax$ subject to $x^T x = 1$,
$A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{pmatrix}$.

For maximum value: compute the characteristic equation of A

$$\det(A - \lambda I) = 0 = (\lambda - 6)(\lambda - 3)(\lambda - 1).$$

Then the maximum value is 6.

For unit vector attaining $Q(x) = 6$: Find eigenvector of A corresponding to 6,

and normalize it!

Get both using a decompostion of A...

Have access to orthogonal diagonalization?

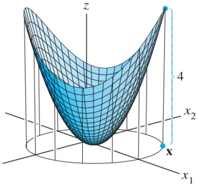
Example

What is the vector attaining the maximum value of $Q(x) = x^T Ax$ subject to $x^T x = 1$, $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$.

If you have the *orthogonal diagonalization of A*:

$$\begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

- The **maximum** value is 4
- and the *vectors attaining* such value are $\pm \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$



The maximum value of $Q(\mathbf{x})$ subject to $\mathbf{x}^T\mathbf{x} = 1$ is 4.