

Econ 204 2022

Lecture 10

Outline

0. Diagonalization

1. Diagonalization of Real Symmetric Matrices
2. Application to Quadratic Forms
3. Linear Maps Between Normed Spaces

Announcements

- PS 3 due today
- PS 4 posted
→ due Tuesday
- last year's exam posted \approx Sunday

- (*) • PS 3 6(a) :
show there
is a finite
or countable
subset $A \subseteq X$
s.t.

$$\lambda \in \mathbb{R} \quad T(v) = \lambda v \quad \Leftrightarrow \quad Ax = \lambda x$$

$$v \neq 0 \quad A = Mtx_U(T), \quad x = \text{crd}_U(v)$$

λ eigenvalue for T , v eigenvector $\Leftrightarrow \lambda$ eigenvalue, x eigenvector for A

Computing Eigenvalues and Eigenvectors

Suppose $\dim X = n$; let I be the $n \times n$ identity matrix. Given $T \in L(X, X)$, fix a basis U and let

$$A = Mtx_U(T)$$

Find the eigenvalues of T by computing the eigenvalues of A :

$$\exists v \neq 0 \text{ s.t. } Av = \lambda v \quad \Leftrightarrow \quad (A - \lambda I)v = 0 \quad \text{for some } v \neq 0$$

$$\Leftrightarrow (A - \lambda I) \text{ is not invertible}$$

$$\Leftrightarrow \det(A - \lambda I) = 0$$

We have the following facts:

$$f(\lambda) = 0$$

- If $A \in \mathbf{R}_{n \times n}$,

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

is an n^{th} degree polynomial in λ with real coefficients; it is called the *characteristic polynomial* of A .

- f has n roots in \mathbf{C} , counting multiplicity:

$$f(\lambda) = (c_1 - \lambda)(c_2 - \lambda) \cdots (c_n - \lambda)$$

(may have
 $c_i = c_j, i \neq j$)

where $c_1, \dots, c_n \in \mathbf{C}$ are the eigenvalues; the c_j 's are not necessarily distinct. Notice that $f(\lambda) = 0$ if and only if $\lambda \in \{c_1, \dots, c_n\}$, so the roots are the solutions of the equation $f(\lambda) = 0$.

- the roots that are not real come in conjugate pairs:

$$f(a + bi) = 0 \Leftrightarrow f(a - bi) = 0$$

- if $\lambda = c_j \in \mathbf{R}$, there is a corresponding eigenvector in \mathbf{R}^n .
- if $\lambda = c_j \notin \mathbf{R}$, the corresponding eigenvectors are in $\mathbf{C}^n \setminus \mathbf{R}^n$.

Diagonalization

Definition 3. Suppose X is a finite-dimensional vector space with basis U . Given a linear transformation $T \in L(X, X)$, let

$$A = Mtx_U(T)$$

We say that A can be diagonalized if there is a basis W for X such that $Mtx_W(T)$ is a diagonal matrix, that is,

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$Mtx_W(T) = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

So

A can be diagonalized $\iff A$ is similar to a diagonal matrix
 $\iff A = P^{-1}BP$ where B is diagonal

u

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \lambda_n \\ 0 & \cdots & \cdots & \lambda_n \end{pmatrix}$$

Suppose there is a basis W such that

$$Mtx_W(T) = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

$\Rightarrow \lambda_1, \dots, \lambda_n$ are eigenvalues of $mtx_W(T)$ and T

Then the standard basis vectors of \mathbf{R}^n are eigenvectors of $Mtx_W(T)$.

In general:

z_j is an eigenvector of T corresponding to $\lambda_j \iff crd_W(z_j)$ is an eigenvector of $Mtx_W(T)$ corresponding to λ_j .

So an eigenvector _{of T} corresponding to λ_j is w_j , since $crd_W(w_j) = e_j$, the j^{th} standard basis vector in \mathbf{R}^n .

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

Thus $Mtx_W(T)$ is diagonal if and only if $W = \{w_1, \dots, w_n\}$ where w_j is an eigenvector of T corresponding to λ_j for each j .

Then the action of T is clear: it stretches each basis element w_i by the factor λ_i .

Diagonalization

Theorem 7 (Thm. 6.7'). *Let X be an n -dimensional vector space, $T \in L(X, X)$, U any basis of X , and $A = Mtx_U(T)$. Then the following are equivalent:*

- 1. A can be diagonalized*
- 2. there is a basis W for X consisting of eigenvectors of T*
- 3. there is a basis V for \mathbf{R}^n consisting of eigenvectors of A*

Proof. Follows from Theorem 6.7 in de la Fuente and Theorem 4 from the Handout. □

Diagonalization

Theorem 8 (Thm. 6.8'). *Let X be a vector space and $T \in L(X, X)$.*

- 1. If $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T with corresponding eigenvectors v_1, \dots, v_m , then $\{v_1, \dots, v_m\}$ is linearly independent.*
- 2. If $\dim X = n$ and T has n distinct eigenvalues, then X has a basis consisting of eigenvectors of T ; consequently, if U is any basis of X , then $Mtx_U(T)$ is diagonalizable.*

Proof. This is an adaptation of the proof of Theorem 6.8 in de la Fuente. □

How Might This Matter

$$\begin{aligned} c_{t+1} &= b_{11}c_t + b_{12}k_t \\ k_{t+1} &= b_{21}c_t + b_{22}k_t \end{aligned}$$

- Why does diagonalizability matter?

Consider a two-dimensional linear difference equation:

$$\begin{pmatrix} c_{t+1} \\ k_{t+1} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} c_t \\ k_t \end{pmatrix} \quad \forall t = 0, 1, 2, 3, \dots$$

given an initial condition c_0, k_0 , or, setting

$$y_t = \begin{pmatrix} c_t \\ k_t \end{pmatrix} \quad \forall t \quad \text{and} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

we can rewrite this more compactly as

$$y_{t+1} = By_t \quad \forall t$$

where $b_{ij} \in \mathbf{R}$ each i, j .

We want to find a solution y_t , $t = 1, 2, 3, \dots$ given initial condition y_0 . (Why?)

Such a dynamical system will arise for example as a characterization of the solution to a standard infinite-horizon optimal growth problem (202a, lecture 2).

If B is diagonalizable, this can be easily solved after a change of basis. If B is diagonalizable, choose an invertible 2×2 real matrix P such that

$$P^{-1}BP = D = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$$

Then

$$\begin{aligned} y_{t+1} = By_t \quad \forall t &\iff P^{-1}y_{t+1} = P^{-1}By_t \quad \forall t && \text{(mult. by } P^{-1}) \\ &\iff P^{-1}y_{t+1} = (P^{-1}BP)(P^{-1}y_t) \quad \forall t && PP^{-1} = I \\ &\iff \bar{y}_{t+1} = D\bar{y}_t \quad \forall t \\ & && = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \bar{y}_t, \quad \bar{y}_t = P^{-1}y_t \quad \forall t \end{aligned}$$

$$\Leftrightarrow \bar{y}_{i,t+1} = d_i \bar{y}_{i,t} \quad \forall t, \quad i=1,2$$

where $\bar{y}_t = P^{-1}y_t \quad \forall t$.

Since D is diagonal, after a change of basis to \bar{y}_t , we need to solve two **independent** linear univariate difference equations, which is easy:

$$\bar{y}_{it} = d_i^t \bar{y}_{i0} \quad \forall t$$

- Not all real $n \times n$ matrices are diagonalizable (not even all invertible $n \times n$ matrices are)...so can we identify some classes that are?
 - basis of eigenvectors (\Leftrightarrow)
 - n distinct eigenvalues (\Rightarrow)
- Some types of matrices appear more frequently than others – especially real symmetric $n \times n$ matrices (matrix representation of second derivatives of C^2 functions, quadratic forms...).
 e.g. second order conditions in optimization, checking concavity and convexity, Taylor series approximation of function

- Recall that an $n \times n$ real matrix A is *symmetric* if $a_{ij} = a_{ji}$ for all i, j , where a_{ij} is the $(i, j)^{\text{th}}$ entry of A .

Rest of this section: work in \mathbb{R}^n

- vector space
- norm
- inner product ($x \cdot y = \sum_{i=1}^n x_i y_i$)

Orthonormal Bases

Definition 1. Let

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

A basis $V = \{v_1, \dots, v_n\}$ of \mathbb{R}^n is orthonormal if $v_i \cdot v_j = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$

In other words, a basis is orthonormal if each basis element has unit length ($\|v_i\|^2 = v_i \cdot v_i = 1 \ \forall i$), and distinct basis elements are perpendicular ($v_i \cdot v_j = 0$ for $i \neq j$).

$$\|x\| = \left(\sum_{i=1}^n x_i^2 \right)^{1/2} = (x \cdot x)^{1/2}$$

Orthonormal Bases

Remark: Suppose that $x = \sum_{j=1}^n \alpha_j v_j$ where $\{v_1, \dots, v_n\}$ is an orthonormal basis of \mathbf{R}^n . Then

$$\begin{aligned} x \cdot v_k &= \left(\sum_{j=1}^n \alpha_j v_j \right) \cdot v_k \\ &= \sum_{j=1}^n \alpha_j (v_j \cdot v_k) \\ &= \sum_{j=1}^n \alpha_j \delta_{jk} = \begin{cases} 1 & j=k \\ 0 & j \neq k \end{cases} \\ &= \alpha_k \end{aligned}$$

so

$$x = \sum_{j=1}^n (x \cdot v_j) v_j$$

Orthonormal Bases

Example: The standard basis of \mathbf{R}^n is orthonormal.

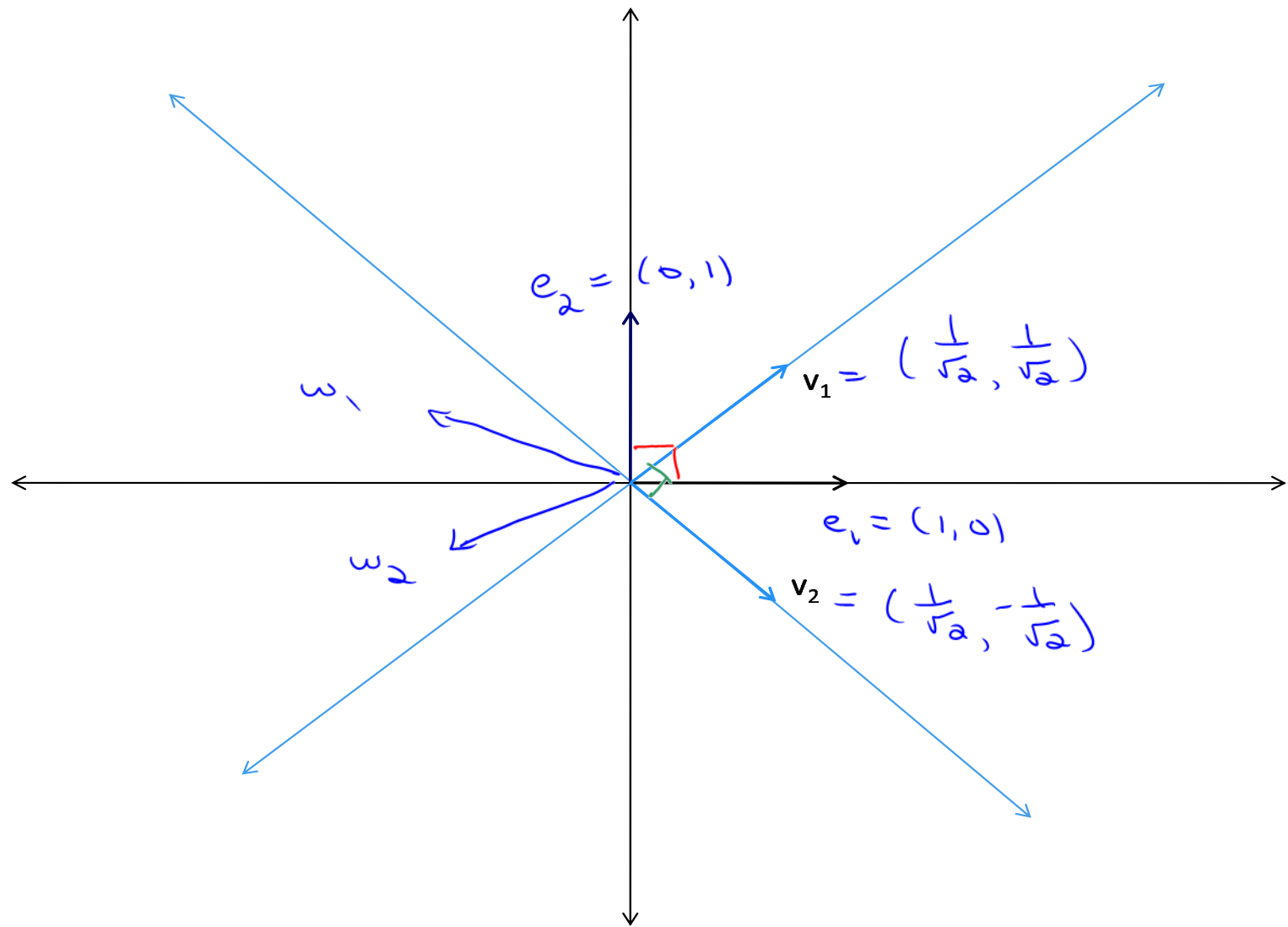
$$e_i = (0, \dots, 1, 0, \dots, 0) \quad i = 1, \dots, n$$

(Why?)

$$\text{e.g. } \mathbb{R}^2 : e_1 = (1, 0), e_2 = (0, 1)$$

$$\text{others? e.g. } v_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), v_2 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$$

also many bases that are not orthonormal



Unitary Matrices

Recall that for a real $n \times m$ matrix A , A^\top denotes the transpose of A : the $(i, j)^{th}$ entry of A^\top is the $(j, i)^{th}$ entry of A .

So the i^{th} row of A^\top is the i^{th} column of A .

Definition 2. A real $n \times n$ matrix A is unitary if $A^\top = A^{-1}$.

$$A^\top A = I$$

Notice that by definition every unitary matrix is invertible.

Unitary Matrices

Theorem 1. *A real $n \times n$ matrix A is unitary if and only if the columns of A are orthonormal.*

Proof. Let v_j denote the j^{th} column of A .

$$\begin{aligned}
 A^T &= A^{-1} &\iff A^T A &= I = (\delta_{ij}) \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \\
 &&\iff v_i \cdot v_j &= \delta_{ij} \quad \forall i, j \\
 &&\iff \{v_1, \dots, v_n\} &\text{ is orthonormal}
 \end{aligned}$$

□

$$A^T = A^{-1}$$

$$A = (\underset{\uparrow}{v_1}, \dots, \underset{\uparrow}{v_n}) \quad - \quad \text{mtx}_{w,v}(\text{id})$$

$$\text{ord}_w(v_1) \quad \text{ord}_w(v_n)$$

Unitary Matrices

$$= \{v_1, \dots, v_n\}$$

If A is unitary, let V be the set of columns of A and W be the standard basis of \mathbf{R}^n . Since A is unitary, it is invertible, so V is a basis of \mathbf{R}^n . ($\{v_1, \dots, v_n\}$ linearly independent)

$$A^T = A^{-1} = \text{mtx}_{V,W}(\text{id}) \quad - \quad \begin{array}{l} \text{change of basis} \\ \text{from } w \text{ to } v \\ \uparrow \\ \text{standard basis} \end{array}$$

Since V is orthonormal, the transformation between bases W and V preserves all geometry, including lengths and angles.

Thus : Let C be an $n \times n$ real symmetric matrix.
Then C is diagonalizable. In addition,

$$C = P^{-1} D P$$

where D is a diagonal matrix and P is unitary.

Note : The diagonal elements $\{\lambda_1, \dots, \lambda_n\}$ of D
are the eigenvalues of C

- C has orthonormal eigenvectors $\{v_1, \dots, v_n\}$
that are a basis for \mathbb{R}^n .

Diagonalization of Real Symmetric Matrices

Theorem 2. *Let $T \in L(\mathbf{R}^n, \mathbf{R}^n)$ and W be the standard basis of \mathbf{R}^n . Suppose that $Mtx_W(T)$ is symmetric. Then the eigenvectors of T are all real, and there is an orthonormal basis $V = \{v_1, \dots, v_n\}$ of \mathbf{R}^n consisting of eigenvectors of T , so that $Mtx_W(T)$ is diagonalizable:*

$$Mtx_W(T) = Mtx_{W,V}(id) \cdot Mtx_V(T) \cdot Mtx_{V,W}(id)$$

where $Mtx_V T$ is diagonal and the change of basis matrices $Mtx_{V,W}(id)$ and $Mtx_{W,V}(id)$ are unitary.

The proof of the theorem requires a lengthy digression into the linear algebra of *complex* vector spaces. A brief outline is in the notes.

quadratic form : polynomial with all terms of degree 2

Quadratic Forms

Example: Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}$

$$f(x) = \alpha x_1^2 + \beta x_1 x_2 + \gamma x_2^2$$

Let

write as $f(x) = x^T A x$, A symmetric

$$A = \begin{pmatrix} \alpha & \frac{\beta}{2} \\ \frac{\beta}{2} & \gamma \end{pmatrix}$$

$$x^T A x = (x_1 \ x_2) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

so A is symmetric and

$$\begin{aligned}x^\top Ax &= (x_1, x_2) \begin{pmatrix} \alpha & \frac{\beta}{2} \\ \frac{\beta}{2} & \gamma \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\&= (x_1, x_2) \begin{pmatrix} \alpha x_1 + \frac{\beta}{2} x_2 \\ \frac{\beta}{2} x_1 + \gamma x_2 \end{pmatrix} \\&= \alpha x_1^2 + \beta x_1 x_2 + \gamma x_2^2 \\&= f(x)\end{aligned}$$

Notice $f(0) = 0$.

Can we determine anything about $f(x)$ for $x \neq 0$?

e.g. $f(x) \geq 0 \quad \forall x$?

easy if $\beta = 0 \dots$

Quadratic Forms

general form:

Consider a quadratic form

$$f(x_1, \dots, x_n) = \sum_{i=1}^n \alpha_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j \quad (1)$$

Let

$$\alpha_{ij} = \begin{cases} \frac{\beta_{ij}}{2} & \text{if } i < j \\ \frac{\beta_{ji}}{2} & \text{if } i > j \end{cases}$$

above diagonal
below diagonal

Let

$$A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{pmatrix} \quad \text{so } f(x) = x^T A x$$

↑
real symmetric

Quadratic Forms

A is symmetric, so let $V = \{v_1, \dots, v_n\}$ be an orthonormal basis of eigenvectors of A with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$.

$$\text{Then } A = U^\top D U = u^{-1} D u$$

$$\text{where } D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

$$u^{-1} \text{ and } U = \text{Mtx}_{V,W}(\text{id}) \text{ is unitary}$$

The columns of U^\top (the rows of U) are the coordinates of v_1, \dots, v_n , expressed in terms of the standard basis W . Given $x \in \mathbf{R}^n$, recall

$$x = \sum_{i=1}^n \gamma_i v_i \text{ where } \gamma_i = x \cdot v_i$$

Quadratic Forms

So

$$\begin{aligned}
 f(x) &= f(x) = f\left(\sum \gamma_i v_i\right) \\
 &= \left(\sum \gamma_i v_i\right)^\top A \left(\sum \gamma_i v_i\right) = x^\top A x \\
 &= \left(\sum \gamma_i v_i\right)^\top U^\top D U \left(\sum \gamma_i v_i\right) \\
 &= \left(U \sum \gamma_i v_i\right)^\top D \left(U \sum \gamma_i v_i\right) \\
 &= \left(\sum \gamma_i U v_i\right)^\top D \left(\sum \gamma_i U v_i\right) \\
 &= (\gamma_1, \dots, \gamma_n) D \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} \\
 &= \sum \lambda_i \gamma_i^2
 \end{aligned}$$

$$(EF)^\top = F^\top E^\top$$

(u linear)

← U is change of basis matrix from W to V
 $\Rightarrow U v_i = e_i \quad \forall i$

↑
 eigenvalues
 of A

Quadratic Forms

$$f(x) = x^T A x$$

This proves the following corollary of Theorem 2.

Corollary 1. Consider the quadratic form (1). Let $\{v_1, \dots, v_n\}$ be an orthonormal basis of eigenvectors of A with corresponding eigenvalues $\{\lambda_1, \dots, \lambda_n\}$

1. f has a global minimum at 0 if and only if $\lambda_i \geq 0$ for all i ; the level sets of f are ellipsoids with principal axes aligned with the orthonormal eigenvectors v_1, \dots, v_n .
2. f has a global maximum at 0 if and only if $\lambda_i \leq 0$ for all i ; the level sets of f are ellipsoids with principal axes aligned with the orthonormal eigenvectors v_1, \dots, v_n .

3. If $\lambda_i < 0$ for some i and $\lambda_j > 0$ for some j , then f has a saddle point at 0 ; the level sets of f are hyperboloids with principal axes aligned with the orthonormal eigenvectors v_1, \dots, v_n .

Bounded Linear Maps

(over \mathbb{R})

Definition 3. Suppose X, Y are normed vector spaces and $T \in L(X, Y)$. We say T is bounded if

$$\exists \beta \in \mathbb{R} \text{ s.t. } \|T(x)\|_Y \leq \beta \|x\|_X \quad \forall x \in X$$

Note this implies that T is Lipschitz with constant β .

why not

$$\exists C \in \mathbb{R} \text{ s.t. } \|T(x)\| \leq C \quad \forall x \in X \quad ?$$

$$T(\alpha x) = \alpha T(x) \quad \forall \alpha \in \mathbb{R} \quad T \text{ linear}$$

$$\Rightarrow \|T(\alpha x)\| = \|\alpha T(x)\| = |\alpha| \|T(x)\| \quad \forall \alpha \in \mathbb{R}$$

Bounded Linear Maps

Much more is true:

Theorem 3 (Thms. 4.1, 4.3). *Let X and Y be normed vector spaces and $T \in L(X, Y)$. Then*

- T is continuous at some point $x_0 \in X$*
- \iff *T is continuous at every $x \in X$*
- \iff *T is uniformly continuous on X*
- \iff *T is Lipschitz*
- \iff *T is bounded*

Proof. Suppose T is continuous at x_0 . Fix $\varepsilon > 0$. Then there exists $\delta > 0$ such that

$$\|z - x_0\| < \delta \Rightarrow \|T(z) - T(x_0)\| < \varepsilon$$

Now suppose x is any element of X . If $\|y - x\| < \delta$, let $z =$
 $z = y - x + x_0$, so $\|z - x_0\| = \|y - x\| < \delta$.

$$\underbrace{z - x_0}_{z - x_0 = y - x} = y - x$$

$$\begin{aligned} & \|T(y) - T(x)\| \\ &= \|T(y - x)\| && (T \text{ linear}) \\ &= \|T(y - x + x_0 - x_0)\| = \|T(z - x_0)\| \\ &= \|T(z) - T(x_0)\| \\ &< \varepsilon \end{aligned}$$

which proves that T is continuous at every x , and uniformly continuous.

We claim that T is bounded if and only if T is continuous at 0. Suppose T is not bounded. Then

$$\exists \{x_n\} \text{ s.t. } \|T(x_n)\| > n\|x_n\| \quad \forall n$$

x_n

$$\|x_n - 0\| =$$

$$\begin{aligned} \|T(x'_n) - T(0)\| &= \|T(x'_n)\| \\ &= \frac{1}{n\|x_n\|} \|T(x_n)\| \\ &> \frac{n\|x_n\|}{n\|x_n\|} \\ &= 1 \\ &= \varepsilon \end{aligned}$$

(defn of x'_n + T linear)

(defn of x_n)

Since this is true for every δ , T is not continuous at 0. Therefore, T continuous at 0 implies T is bounded. Now, suppose T is bounded, so find M such that $\|T(x)\| \leq M\|x\|$ for every $x \in X$. Given $\varepsilon > 0$, let $\delta = \varepsilon/M$. Then

$$\begin{aligned} \|x - 0\| < \delta &\Rightarrow \|x\| < \delta \\ &\Rightarrow \|T(x) - T(0)\| = \|T(x)\| < M\delta \quad (\text{defn of } M) \\ &\Rightarrow \|T(x) - T(0)\| < \varepsilon = M\delta \end{aligned}$$

so T is continuous at 0.

Thus, we have shown that continuity at some point x_0 implies uniform continuity, which implies continuity at every point, which implies T is continuous at 0, which implies that T is bounded, which implies that T is continuous at 0, which implies that T is

continuous at some x_0 , so all of the statements except possibly the Lipschitz statement are equivalent.

Suppose T is bounded, with constant M ^{>0}. Then

$$\begin{aligned} \|T(x) - T(y)\| &= \|T(x - y)\| && (T \text{ linear}) \\ &\leq M\|x - y\| \end{aligned}$$

so T is Lipschitz with constant M ; conversely, if T is Lipschitz with constant M , then T is bounded with constant M . So all the statements are equivalent. \square

$$\forall x \in X: \|T(x) - T(0)\| = \|T(x)\| \leq M\|x - 0\| = M\|x\|$$

Bounded Linear Maps

Every linear map on a finite-dimensional normed vector space is bounded (and thus continuous, uniformly continuous, and Lipschitz continuous).

Theorem 4 (Thm. 4.5). *Let X and Y be normed vector spaces, with $\dim X = n$. Every $T \in L(X, Y)$ is bounded.*

($n \in \mathbb{N}$)

Proof. See de la Fuente.



Topological Isomorphism

Definition 4. A topological isomorphism *between normed vector spaces X and Y* is a linear transformation $T \in L(X, Y)$ that is invertible (one-to-one, onto), continuous, and has a continuous inverse.

Two normed vector spaces X and Y are topologically isomorphic if there is a topological isomorphism $T : X \rightarrow Y$.

The Space $B(X, Y)$

T bounded :
 $\exists \beta > 0$ s.t.
 $\|T(x)\| \leq \beta \|x\| \quad \forall x$

$$\Rightarrow \frac{\|T(x)\|}{\|x\|} \leq \beta \quad \forall x \neq 0$$

Suppose X and Y are normed vector spaces. We define

$$B(X, Y) = \{T \in L(X, Y) : T \text{ is bounded}\}$$

Define :

$$\begin{aligned} \|T\|_{B(X, Y)} &= \sup \left\{ \frac{\|T(x)\|_Y}{\|x\|_X} : x \in X, x \neq 0 \right\} \\ &= \sup \{ \|T(x)\|_Y : \|x\|_X = 1 \} \end{aligned}$$

* note : $\Rightarrow \|T(x)\| \leq \|T\| \|x\| \quad \forall x \in X$ by definition

We skip the proofs of the rest of these results – read d1F.

$$\frac{\|T(x)\|}{\|x\|} \leq \|T\| \quad \forall x \neq 0$$

The Space $B(X, Y)$

Theorem 5 (Thm. 4.8). *Let X, Y be normed vector spaces. Then*

$$\left(B(X, Y), \|\cdot\|_{B(X, Y)} \right)$$

is a normed vector space.

representation

The Space $B(\mathbb{R}^n, \mathbb{R}^m)$

Theorem 6 (Thm. 4.9). Let $T \in L(\mathbb{R}^n, \mathbb{R}^m)$ ($= B(\mathbb{R}^n, \mathbb{R}^m)$) with matrix $A = (a_{ij})$ with respect to the standard bases. Let

$$M = \max\{|a_{ij}| : 1 \leq i \leq m, 1 \leq j \leq n\}$$

Then

$$M \leq \|T\| \leq M\sqrt{mn}$$

.

Compositions

Theorem 7 (Thm. 4.10). *Let $R \in L(\mathbf{R}^m, \mathbf{R}^n)$ and $S \in L(\mathbf{R}^n, \mathbf{R}^p)$.
Then*

$$\|S \circ R\| \leq \|S\| \|R\|$$

Invertibility

Define $\Omega(\mathbf{R}^n) = \{T \in L(\mathbf{R}^n, \mathbf{R}^n) : T \text{ is invertible}\}$

Theorem 8 (Thm. 4.11'). *Suppose $T \in L(\mathbf{R}^n, \mathbf{R}^n)$ and E is the standard basis of \mathbf{R}^n . Then*

T is invertible

$$\iff \ker T = \{0\}$$

$$\iff \det(Mtx_E(T)) \neq 0$$

$$\iff \det(Mtx_{V,V}(T)) \neq 0 \text{ for every basis } V$$

$$\iff \det(Mtx_{V,W}(T)) \neq 0 \text{ for every pair of bases } V, W$$

} nice exercise

Invertibility

Theorem 9 (Thm. 4.12). *If $S, T \in \Omega(\mathbf{R}^n)$, then $S \circ T \in \Omega(\mathbf{R}^n)$ and*

$$(S \circ T)^{-1} = T^{-1} \circ S^{-1}$$

Invertibility

Theorem 10 (Thm. 4.14). *Let $S, T \in L(\mathbf{R}^n, \mathbf{R}^n)$. If T is invertible and*

$$\|T - S\| < \frac{1}{\|T^{-1}\|}$$

then S is invertible. In particular, $\Omega(\mathbf{R}^n)$ is open in $L(\mathbf{R}^n, \mathbf{R}^n) = B(\mathbf{R}^n, \mathbf{R}^n)$.

Theorem 11 (Thm. 4.15). *The function $(\cdot)^{-1} : \Omega(\mathbf{R}^n) \rightarrow \Omega(\mathbf{R}^n)$ that assigns T^{-1} to each $T \in \Omega(\mathbf{R}^n)$ is continuous.*

Quadratic Forms

The equation for a level set of f is

$$\{\gamma \in \mathbb{R}^n : f(\gamma) = C\} = \left\{ \gamma \in \mathbb{R}^n : \sum_{i=1}^n \lambda_i \gamma_i^2 = C \right\} \quad C \in \mathbb{R}$$

- If $\lambda_i \geq 0$ for all i , the level set is an ellipsoid, with principal axes in the directions v_1, \dots, v_n . The length of the principal axis along v_i is $\sqrt{C/\lambda_i}$ if $C \geq 0$ (if $\lambda_i = 0$, the level set is a degenerate ellipsoid with principal axis of infinite length in that direction). The level set is empty if $C < 0$.

$\Rightarrow f$ has global min at 0, $f(x) \geq 0 \quad \forall x$

- If $\lambda_i \leq 0$ for all i , the level set is an ellipsoid, with principal axes in the directions v_1, \dots, v_n . The length of the principal

$\Rightarrow f$ has global max at 0, $f(x) \leq 0 \quad \forall x$

axis along v_i is $\sqrt{C/\lambda_i}$ if $C \leq 0$ (if $\lambda_i = 0$, the level set is a degenerate ellipsoid with principal axis of infinite length in that direction). The level set is empty if $C > 0$.

- If $\lambda_i > 0$ for some i and $\lambda_j < 0$ for some j , the level set is a hyperboloid. For example, suppose $n = 2$, $\lambda_1 > 0$, $\lambda_2 < 0$. The equation is

$$\begin{aligned} C &= \lambda_1 \gamma_1^2 + \lambda_2 \gamma_2^2 \\ &= \left(\sqrt{\lambda_1} \gamma_1 + \sqrt{|\lambda_2|} \gamma_2 \right) \left(\sqrt{\lambda_1} \gamma_1 - \sqrt{|\lambda_2|} \gamma_2 \right) \end{aligned}$$

$\Rightarrow f$ has a saddle point at 0
min with respect to v_i
max with respect to v_j

This is a hyperbola with asymptotes

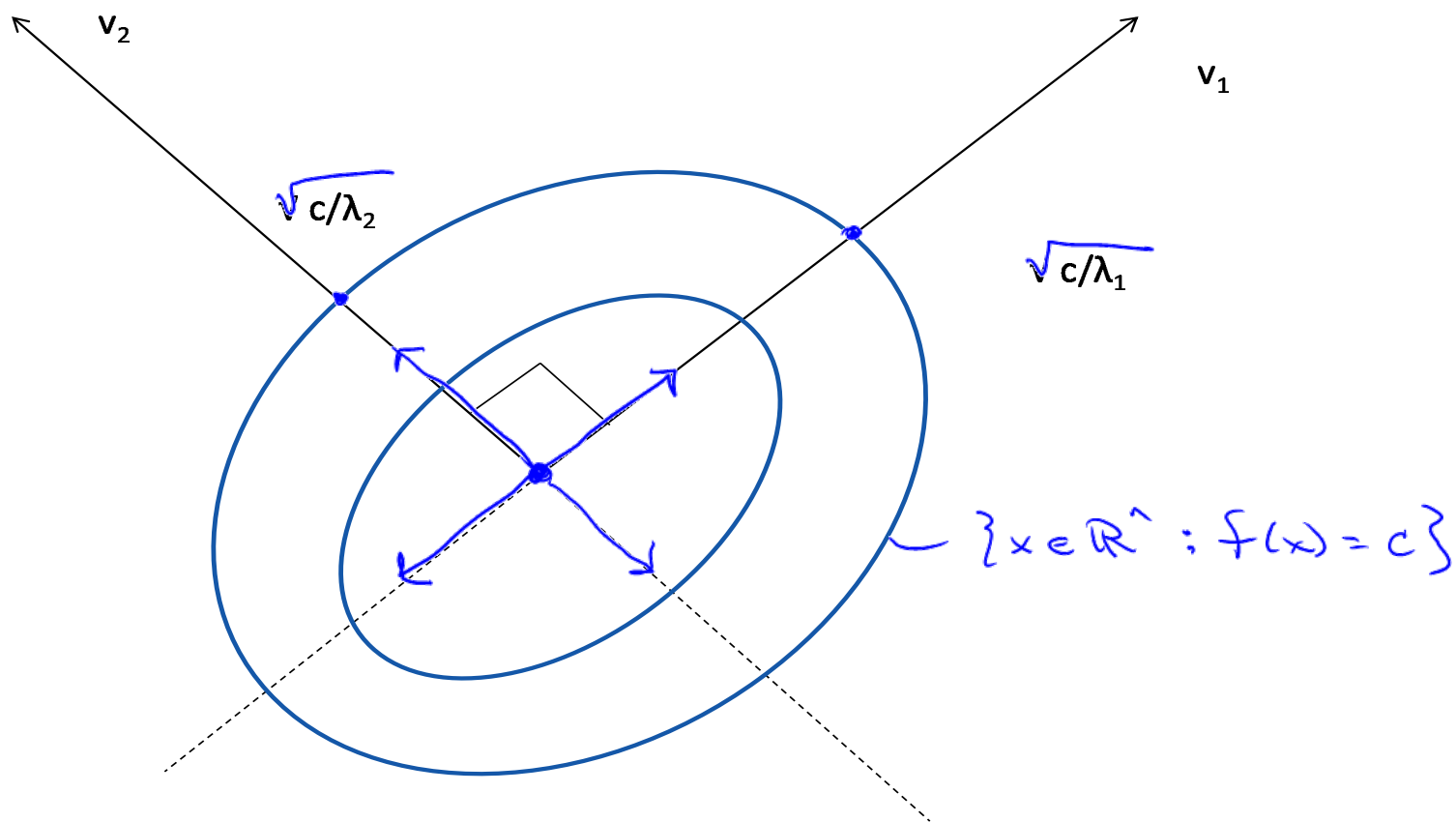
$$0 = \sqrt{\lambda_1}\gamma_1 + \sqrt{|\lambda_2|}\gamma_2$$

$$\Rightarrow \gamma_1 = -\sqrt{\frac{|\lambda_2|}{\lambda_1}}\gamma_2$$

$$0 = \left(\sqrt{\lambda_1}\gamma_1 - \sqrt{|\lambda_2|}\gamma_2 \right)$$

$$\Rightarrow \gamma_1 = \sqrt{\frac{|\lambda_2|}{\lambda_1}}\gamma_2$$

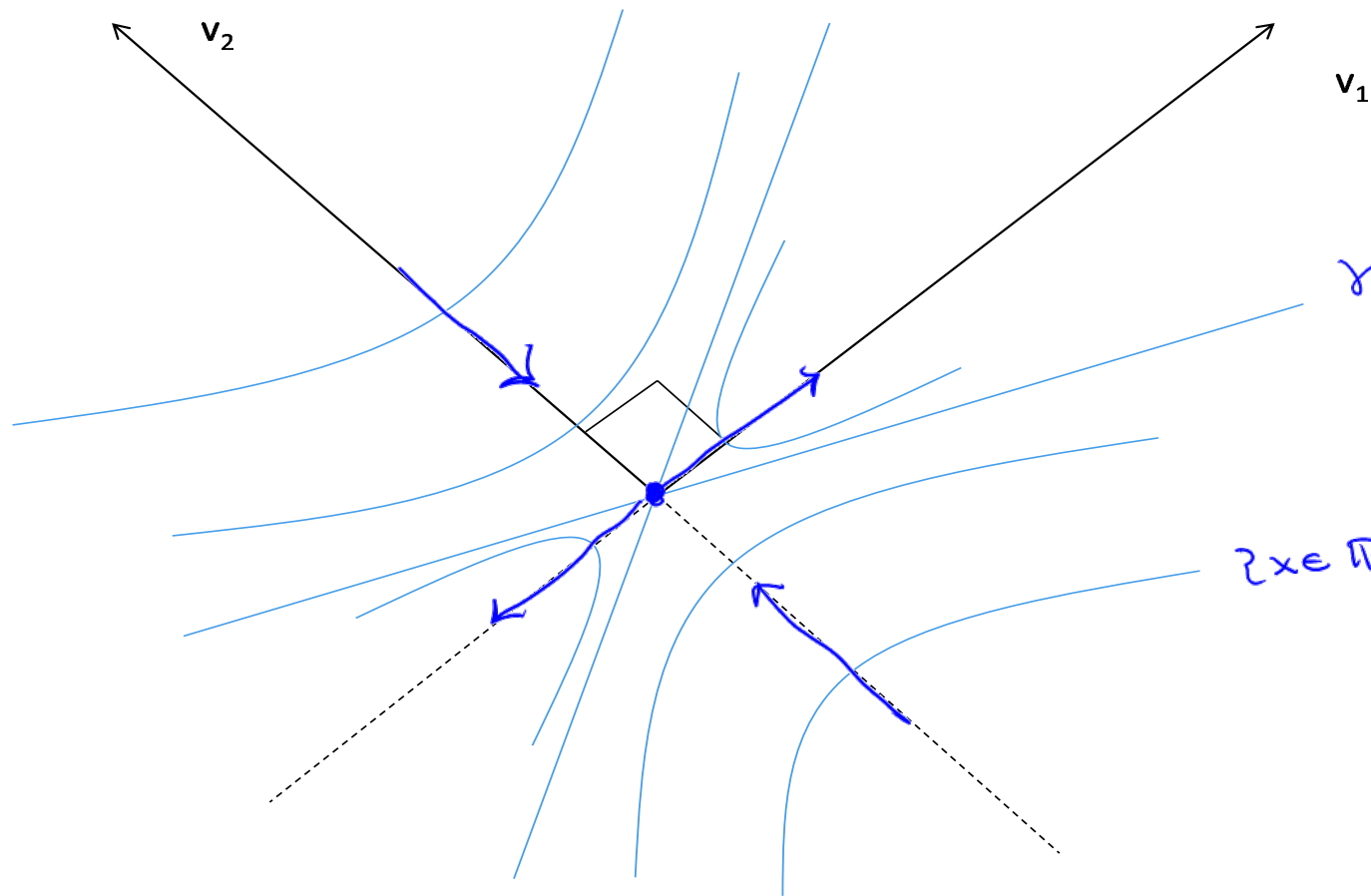
$$\lambda_1 > 0, \lambda_2 > 0$$



f has a global min at 0

$$\lambda_1 > 0, \lambda_2 < 0$$

$$v_1 = \sqrt{|\lambda_2|/\lambda_1} \gamma_2$$



$$\gamma_1 = -\sqrt{\frac{|\lambda_2|}{\lambda_1}} \gamma_2$$

$$\{x \in \mathbb{R}^n : f(x) = c\}$$

f has a saddle point at 0

