Econ 204 2016

Lecture 10

Outline

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

How Might This Matter

- Why does diagonalizability matter?

  Consider a two-dimensional linear difference equation:

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

  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 $y_t, \ t = 1, 2, 3, \ldots$ 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

$$\begin{aligned} y_{t+1} = By_t \ \ \forall t \quad & \Longleftrightarrow \quad P^{-1}y_{t+1} = P^{-1}By_t \ \ \forall t \\ & \Longleftrightarrow \quad P^{-1}y_{t+1} = P^{-1}BPP^{-1}y_t \ \ \forall t \\ & \Longleftrightarrow \quad \bar{y}_{t+1} = D\bar{y}_t \ \ \forall t \end{aligned}$$

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

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?

- 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...).

- 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.

Orthonormal Bases

Definition 1. *Let*

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

A basis $V = \{v_1, \ldots, v_n\}$ *of* $\mathbf{R}^n$ *is* orthonormal *if* $v_i \cdot v_j = \delta_{ij}$.

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 $i \neq j$).

Orthonormal Bases

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

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

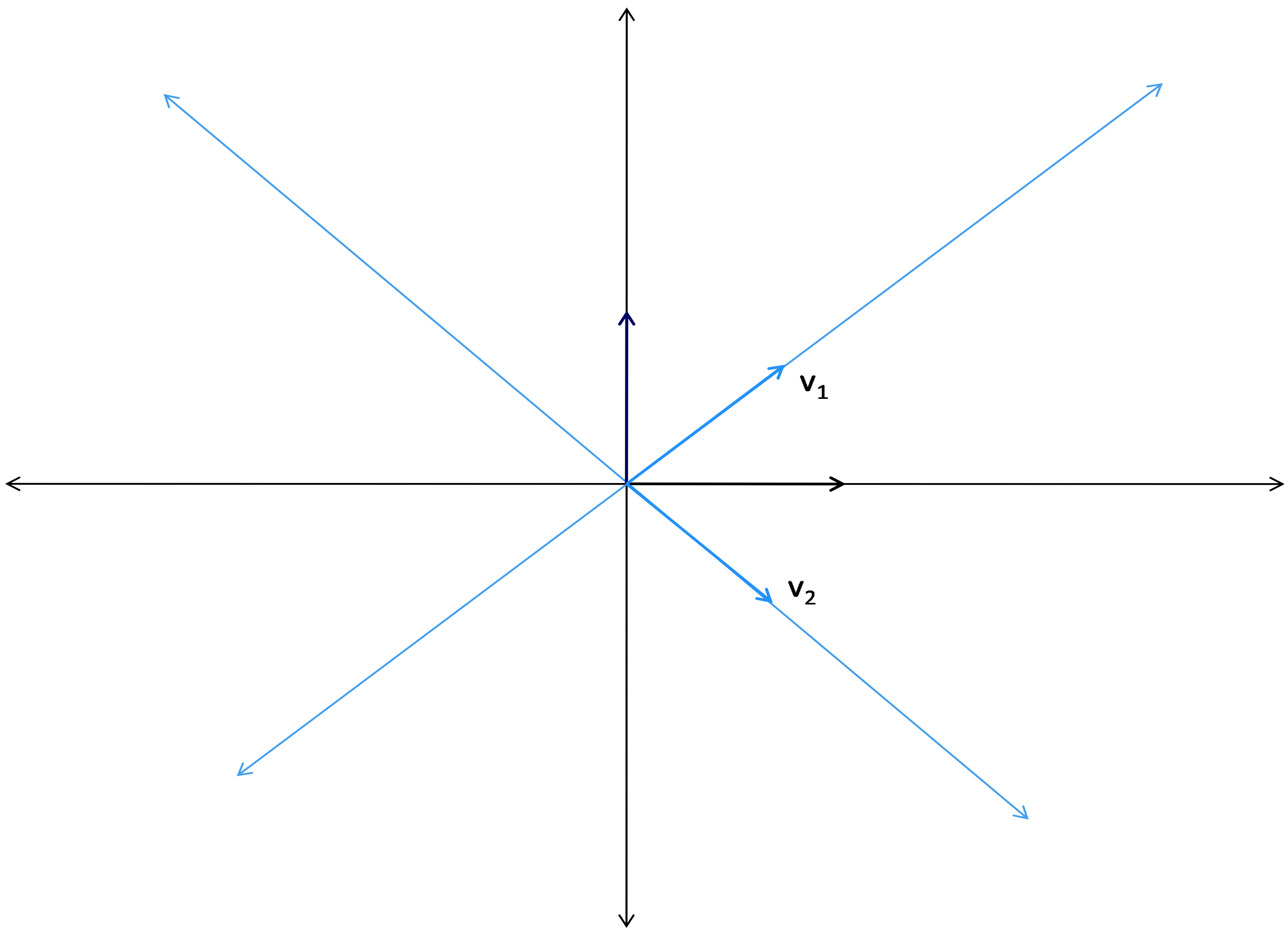
so

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

Orthonormal Bases

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

(Why?)



v_1
v_2

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 $A^\top = A^{-1}$.*

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.

$$\begin{aligned} A^\top = A^{-1} &\iff A^\top A = I \\ &\iff v_i \cdot v_j = \delta_{ij} \quad \forall i, j \\ &\iff \{v_1, \ldots, v_n\} \text{ is orthonormal} \end{aligned}$$

□

Unitary Matrices

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 $\mathbf{R}^n$.

$$A^\top = A^{-1} = Mtx_{V,W}(id)$$

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

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* $V = \{v_1, \ldots, v_n\}$ *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 Forms

Example: Let

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

Let

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

so A is symmetric and

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

Quadratic Forms

Consider a quadratic form

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

Let

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

Let

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

Quadratic Forms

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

$$\begin{aligned} \text{Then } A &= U^\top D U \\ \text{where } D &= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \\ \text{and } U &= Mtx_{V,W}(id) \text{ is unitary} \end{aligned}$$

The columns of $U^\top$ (the rows of U) are the coordinates of $v_1, \ldots, v_n$, 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

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

Quadratic Forms

The equation for a level set of f is

$$\left\{\gamma \in \mathbf{R}^n : \sum_{i=1}^{n} \lambda_i \gamma_i^2 = C\right\}$$

- If $\lambda_i \geq 0$ for all i, the level set is an ellipsoid, with principal axes in the directions $v_1, \ldots, v_n$. 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 $C < 0$.

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

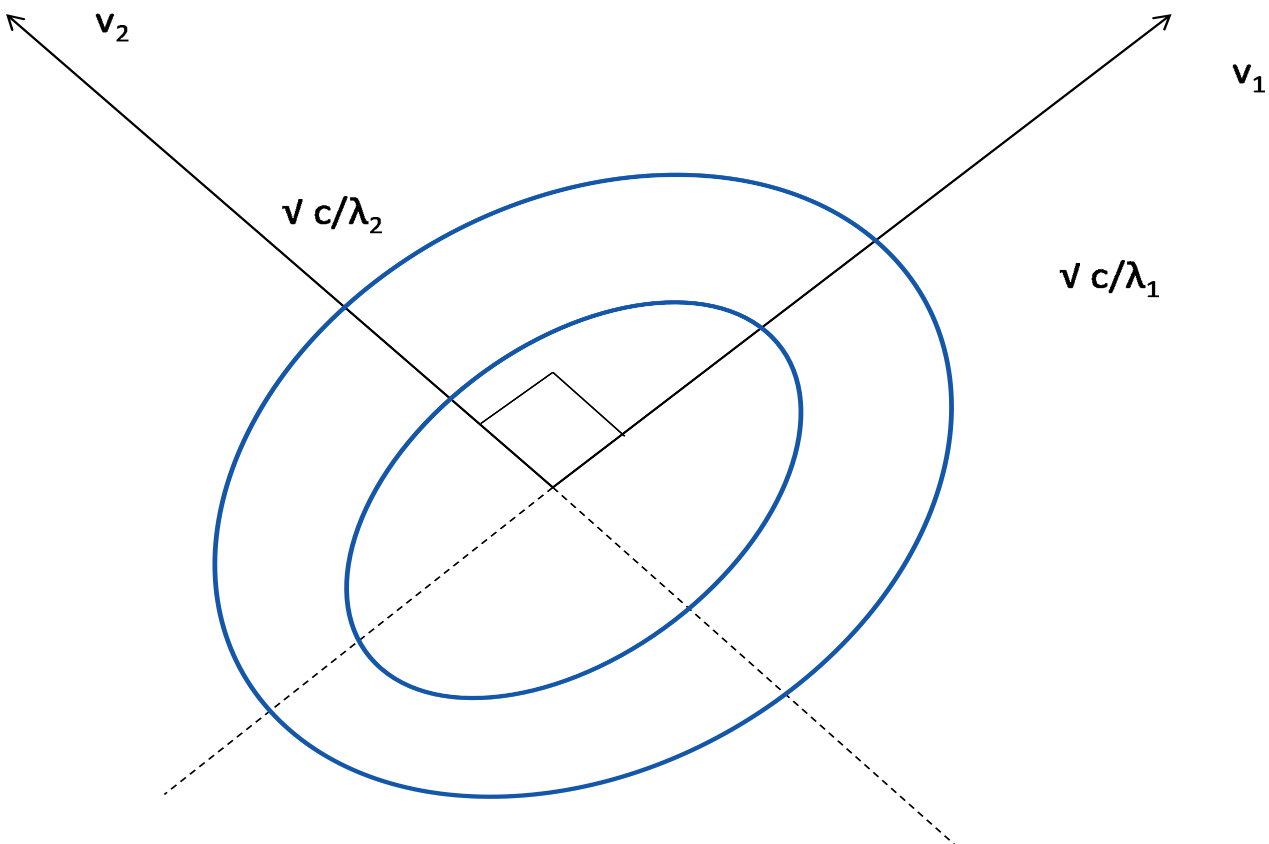
- If $\lambda_i > 0$ for some i and $\lambda_j < 0$ for some j, the level set is a hyperboloid. For example, suppose $n = 2$, $\lambda_1 > 0$, $\lambda_2 < 0$. The equation is

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

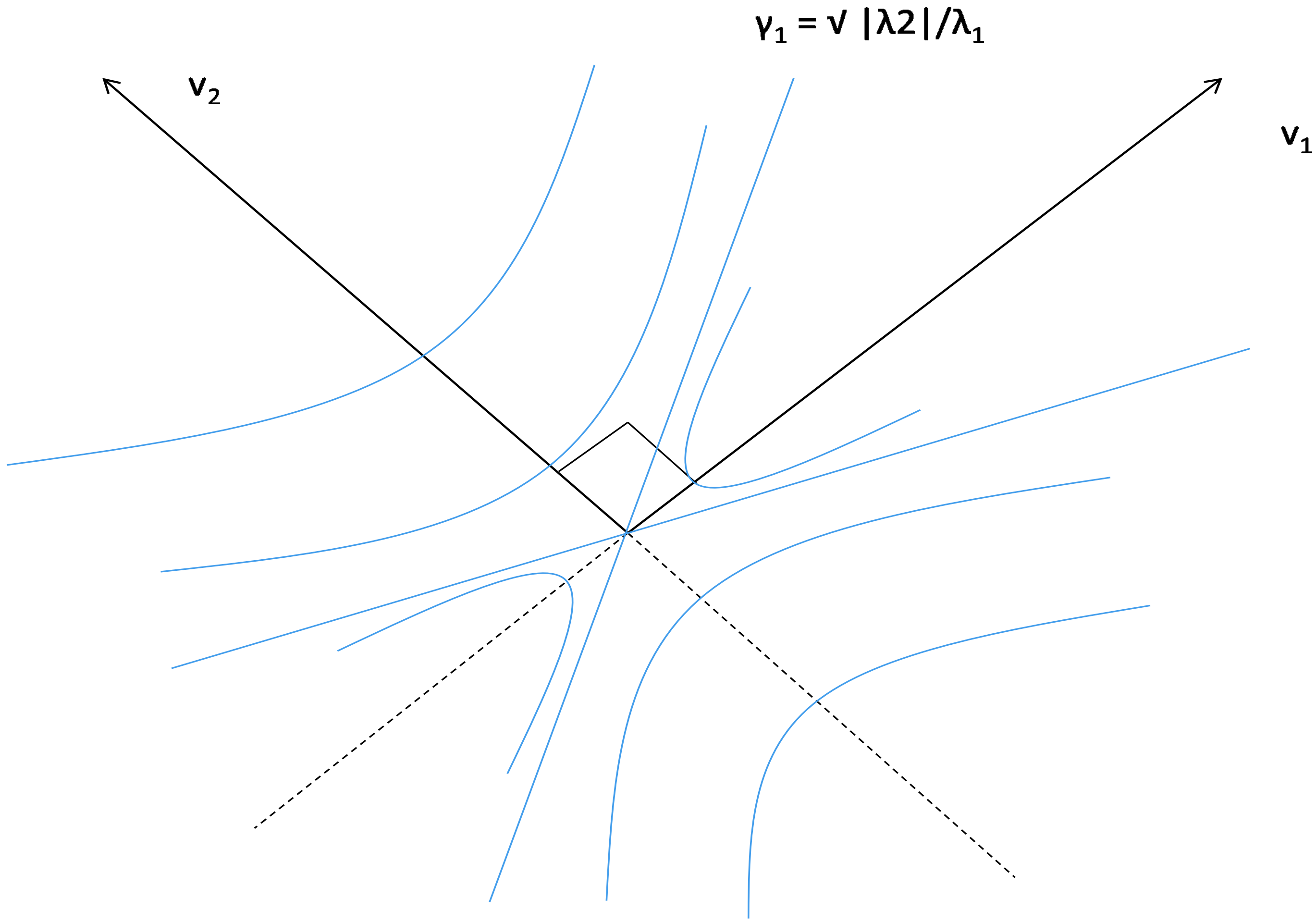
This is a hyperbola with asymptotes

$$
\begin{aligned}
0 &= \sqrt{\lambda_1}\gamma_1 + \sqrt{|\lambda_2|}\gamma_2 \\
\Rightarrow \gamma_1 &= -\sqrt{\frac{|\lambda_2|}{\lambda_1}}\gamma_2 \\
0 &= \left(\sqrt{\lambda_1}\gamma_1 - \sqrt{|\lambda_2|}\gamma_2\right) \\
\Rightarrow \gamma_1 &= \sqrt{\frac{|\lambda_2|}{\lambda_1}}\gamma_2
\end{aligned}
$$

$\lambda_1 > 0, \lambda_2 > 0$



$\lambda_1 > 0, \lambda_2 < 0$



Quadratic Forms

This proves the following corollary of Theorem 2.

Corollary 1. *Consider the quadratic form (1).*

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* $v_1, \ldots, v_n$*.*

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* $v_1, \ldots, v_n$*.*

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 $v_1, \ldots, v_n$.

Bounded Linear Maps

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

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

Note this implies that T is Lipschitz with constant β.

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*

$$
\begin{aligned}
& T \text{ is continuous at some point } x_0 \in X \\
\iff \quad & T \text{ is continuous at every } x \in X \\
\iff \quad & T \text{ is uniformly continuous on } X \\
\iff \quad & T \text{ is Lipschitz} \\
\iff \quad & T \text{ is bounded}
\end{aligned}
$$

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 $z = y - x + x_0$, so $\|z - x_0\| = \|y - x\| < \delta$.

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

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

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

Note that $x_n \neq 0$. Let $\varepsilon = 1$. Fix $\delta > 0$ and choose n such that $\frac{1}{n} < \delta$. Let

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

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

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

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. Then

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

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. □

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.*

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 $T : X \to Y$.*

The Space $B(X, Y)$

Suppose X and Y are normed vector spaces. We define

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

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

The Space $B(X,Y)$

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

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

is a normed vector space.

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

Theorem 6 (Thm. 4.9)**.** *Let* $T \in L(\mathbf{R}^n, \mathbf{R}^m)$ *(*$= B(\mathbf{R}^n, \mathbf{R}^m)$*) 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*

$$
\begin{aligned}
& T \text{ is invertible} \\
\iff & \ker T = \{0\} \\
\iff & \det\left(Mtx_E(T)\right) \neq 0 \\
\iff & \det\left(Mtx_{V,V}(T)\right) \neq 0 \text{ for every basis } V \\
\iff & \det\left(Mtx_{V,W}(T)\right) \neq 0 \text{ for every pair of bases } V, W
\end{aligned}
$$

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 $(\cdot)^{-1} : \Omega(\mathbf{R}^n) \to \Omega(\mathbf{R}^n)$ that assigns T^{-1} to each $T \in \Omega(\mathbf{R}^n)$ is continuous.*