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 \(b_{ij} \in \mathbb{R}\) 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 \( 2 \times 2 \) real matrix \( P \) such that

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

Then

\[
y_{t+1} = By_t \quad \forall t \iff P^{-1}y_{t+1} = P^{-1}By_t \quad \forall t
\]

\[
\iff \begin{pmatrix} P^{-1}y_{t+1} \\ P^{-1}y_t \end{pmatrix} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} P^{-1}y_t \\ P^{-1}y_t \end{pmatrix} \quad \forall t
\]

\[
\iff \bar{y}_{t+1} = D\bar{y}_t \quad \forall t
\]

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

\[
\bar{y}_{it} = d^t_i \bar{y}_0 \quad \forall t
\]

- 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?

  yesterday:  
  - a basis of eigenvectors \( (\Leftarrow) \)
  - \( n \) distinct eigenvalues \( (\Leftarrow) \)

- 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 problems, convexity and concavity checks, ...
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 $(i, j)^{th}$ entry of $A$. 
Orthonormal Bases

Definition 1. Let

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

A basis \( V = \{v_1, \ldots, v_n\} \) of \( \mathbb{R}^n \) 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 \( \mathbb{R}^n \). Then

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

so

\[
x = \sum_{j=1}^{n} (x \cdot v_j) v_j
\]
Orthonormal Bases

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

(Why?)

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

others? $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$

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

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 $A$.

\[
A^\top = A^{-1} \iff A^\top A = I \iff v_i \cdot v_j = \delta_{ij} \quad \forall i, j \\
\iff \{v_1, \ldots, v_n\} \text{ is orthonormal}
\]

\[\square\]
Unitary Matrices

If $A$ is unitary, let $V$ be the set of columns of $A$ and $W$ be the standard basis of $\mathbb{R}^n$. Since $A$ is unitary, it is invertible, so $V$ is a basis of $\mathbb{R}^n$.

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

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 $T \in L(\mathbb{R}^n, \mathbb{R}^n)$ and $W$ be the standard basis of $\mathbb{R}^n$. 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 $\mathbb{R}^n$ consisting of eigenvectors of $T$, so that $Mtx_W(T)$ is diagonalizable:

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

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

i.e. $C = P^{-1}DP$ where $\{\lambda_1, \ldots, \lambda_n\}$

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

\[ f(x) = x^T A x \quad \text{as} \quad A \text{ symmetric} \]

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

\[ x^T A x = (x_1, x_2) \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \]
so $A$ is symmetric and

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

$$= (x_1, x_2) \begin{pmatrix} \alpha x_1 + \frac{\beta}{2} x_2 \\ \frac{\beta}{2} x_1 + \gamma x_2 \end{pmatrix}$$

$$= \alpha x_1^2 + \beta x_1 x_2 + \gamma x_2^2$$

$$= f(x)$$

Notice $f(0) = 0$.

Can we determine anything about $f(x)$ at $x=0$?

E.g., $f(x) > 0$ at $x$? Easy if $\beta = 0$
Quadratic Forms

Consider a quadratic form

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

Let

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

Let

\[ A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{pmatrix} \]

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 $\lambda_1, \ldots, \lambda_n$.

Then $A = U^T D U$

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

and $U = Mtx_{V,W}(id)$ is unitary

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 $x \in \mathbb{R}^n$, recall

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

So

\[ f(x) = f \left( \sum \gamma_i v_i \right) \]

\[ = \left( \sum \gamma_i v_i \right)^T A \left( \sum \gamma_i v_i \right) \]

\[ = \left( \sum \gamma_i v_i \right)^T U^T D U \left( \sum \gamma_i v_i \right) \]

\[ = \left( U \sum \gamma_i v_i \right)^T D \left( U \sum \gamma_i v_i \right) \]

\[ = \left( \sum \gamma_i U v_i \right)^T D \left( \sum \gamma_i U v_i \right) \]

\[ = \left( \gamma_1, \ldots, \gamma_n \right) D \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} \]

\[ = \sum \lambda_i \gamma_i^2 \]

\[ \text{eigenvalues of } A \]

\[ (CEF)^T = F^T E^T \]

\[ U \text{ is change of basis from } W \text{ to } V \Rightarrow \]

\[ U v_i = (0, \ldots, 1, \ldots, 0) \]

\[ \Rightarrow \lambda_i = \text{eigenvector of } v_i \]
Quadratic Forms

The equation for the level sets of $f$ is

$$\sum_{i=1}^{n} \lambda_i \gamma_i^2 = C$$

- 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 $C < 0$.

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

- 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 $C < 0$ (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 max at $0$, $f(x) \leq 0 \forall x$
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

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

$\Rightarrow$ $f$ has a saddle point at $0$ min with respect to $v_i$ max with respect to $v_j$. 
This is a hyperbola with asymptotes

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

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

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

\[
\Rightarrow \gamma_1 = \sqrt{\frac{|\lambda_2|}{\lambda_1}} \gamma_2
\]
\( \lambda_1 > 0, \lambda_2 > 0 \)

\[ \sqrt{c/\lambda_2}, \quad \sqrt{c/\lambda_1} \]

\( f \) has global min at 0
$\lambda_1 > 0, \lambda_2 < 0$

$\gamma_1 = \sqrt{|\lambda_2|}/\lambda_1$

$f$ has a saddle point at $0$
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

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

Note this implies that $T$ is Lipschitz with constant $\beta$. 

Why not

$$\|T(x)\|_Y \leq B \quad \forall x ?$$

$$T(ax) = aT(x) \quad \forall a \in \mathbb{R}$$

$$\Rightarrow \quad \|T(ax)\|_Y = |a| \|T(x)\|_Y \quad \forall a \in \mathbb{R}$$
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 $x_0 \in X$ if and only if $T$ is continuous at every $x \in X$.
- $T$ is uniformly continuous on $X$ if and only if $T$ is Lipschitz.
- $T$ is Lipschitz if and only if $T$ is bounded.

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

\[
(\forall \tilde{z}) \quad \|\tilde{z} - x_0\| < \delta \Rightarrow \|T(\tilde{z}) - T(x_0)\| < \varepsilon
\]
Now suppose $x$ is any element of $X$. If $\|y - x\| < \delta$, let $z = y - x + x_0$, so $\|z - x_0\| = \|y - x\| < \delta$.

\[
\|T(y) - T(x)\|
= \|T(y - x)\| \quad \text{by linearity of } T
\geq \|T(y - x + x_0 - x_0)\|
= \|T(z) - T(x_0)\|
< \varepsilon
\]

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

\[
\exists \{x_n\} \text{ s.t. } \|T(x_n)\| > n\|x_n\| \quad \forall n
\]
Note that $x_n \neq 0$. Let $\varepsilon = 1$. Fix $\delta > 0$ and choose $n$ such that $\frac{1}{n} < \delta$. Let

\[
\begin{align*}
x'_n &= \frac{x_n}{n\|x_n\|} \\
\|x'_n\| &= \frac{\|x_n\|}{n\|x_n\|} \\
&= \frac{1}{n} \\
&< \frac{1}{\delta} \\
\|T(x'_n) - T(0)\| &= \frac{1}{n\|x_n\|} \|T(x_n)\| \\
&> \frac{n\|x_n\|}{n\|x_n\|} \\
&= 1 \\
&= \varepsilon
\end{align*}
\]
Since this is true for every $\delta$, $T$ 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

\[
\|x - 0\| < \delta \Rightarrow \|x\| < \delta \\
\Rightarrow \|T(x) - T(0)\| = \|T(x)\| < M\delta \\
\Rightarrow \|T(x) - T(0)\| < \varepsilon = M\delta
\]

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

$$\|T(x) - T(y)\| = \|T(x - y)\| \leq M\|x - y\|$$

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 $\dim X < \infty$. 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

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

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

$$= \sup\{\|T(x)\|_Y : \|x\|_X = 1\}$$

We skip the proofs of the rest of these results – read dIF.
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.
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 $R \in L(\mathbb{R}^m, \mathbb{R}^n)$ and $S \in L(\mathbb{R}^n, \mathbb{R}^p)$. Then

$$\|S \circ R\| \leq \|S\|\|R\|$$
Define \( \Omega(\mathbb{R}^n) = \{ T \in L(\mathbb{R}^n, \mathbb{R}^n) : T \text{ is invertible} \} \)

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

\[ T \text{ is invertible} \iff \ker T = \{0\} \iff \det (Mtx_E(T)) \neq 0 \iff \det (Mtx_{V,V}(T)) \neq 0 \text{ for every basis } V \iff \det (Mtx_{V,W}(T)) \neq 0 \text{ for every pair of bases } V, W \]
Invertibility

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

\[
(S \circ T)^{-1} = T^{-1} \circ S^{-1}
\]
Invertibility

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

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

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

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