# Econ 204
# Supplement to Section 3.6
# Diagonalization and Quadratic Forms

De La Fuente notes that, if an $n \times n$ matrix has $n$ distinct eigenvalues, it can be diagonalized. In this supplement, we will provide an additional, very important diagonalization result. Symmetric matrices can always be diagonalized; moreover, the change of basis matrices that carry out the diagonalization have a special form. This has an important application to quadratic forms, which in turn have application to the geometry of level sets of preferences, and to the analysis of variance-covariance matrices.

## 1 Diagonalization and Change of Basis

Before proceeding with the diagonalization result for symmetric matrices, it is useful to discuss the relationship between diagonalization and change of basis. De La Fuente (page 151) defines a square matrix $M$ to be *diagonalizable* if there exists an invertible matrix $P$ such that $P^{-1}MP$ is diagonal; he also (page 146) defines two square matrices $A$ and $B$ to be *similar* if there is an invertible matrix $P$ such that $P^{-1}AP = B$, so a square matrix $M$ is diagonalizable if and only if it is similar to a diagonal matrix. Theorem 2 tells us that a matrix is diagonalizable if and only if there is another basis so that the representation of the same transformation in the new basis is diagonal.

**Proposition 1** *Fix an $n$-dimensional vector space $X$ and a basis $U = \{u_1, \ldots, u_n\}$. An $n \times n$ matrix $P$ is invertible if and only if there is a basis $W$ such that $P = (Mtx)_{U,W}(id)$; in this case, $((Mtx)_{U,W}(id))^{-1} = (Mtx)_{W,U}(id)$.*

**Proof:** Suppose that $P$ is invertible. Let $W = \{w_1, \ldots, w_n\}$, where $w_j = \sum_{i=1}^{n} p_{ij} u_i$. Since $P$ is invertible, rank $P = n$, so $W$ is a basis. $P = (Mtx)_{U,W}(id)$.

Conversely, suppose there is a basis $W$ such that $P = (Mtx)_{U,W}(id)$. Then $w_j = \sum_{i=1}^{n} p_{ij} u_i$. By the Commutative Diagram Theorem (see the Supplement to Section 3.3),

$$\begin{aligned}(Mtx)_{U,W}(id) \cdot (Mtx)_{W,U}(id) &= (Mtx)_{U,U}(id \circ id) \\ &= (Mtx)_U(id) \\ &= I\end{aligned}$$

so $P$ is invertible and $(Mtx)_{W,U}(id) = ((Mtx)_{U,W}(id))^{-1}$. ■

**Theorem 2** *Suppose that $X$ is finite-dimensional*

- *If $T \in L(X,X)$, and $U, W$ are any two bases of $X$, then $(Mtx)_W(T)$ and $(Mtx)_U(T)$ are similar.*
- *Conversely, given similar matrices $A, B$ with $A = P^{-1}BP$ and any basis $U$, there is a basis $W$ and $T \in L(X,X)$ such that $B = (Mtx)_U(T)$, $A = (Mtx)_W(T)$, $P = (Mtx)_{U,W}(id)$, and $P^{-1} = (Mtx)_{W,U}(id)$.*

**Proof:** For the first bullet, note that

$$(Mtx)_W(T) = (Mtx)_{W,U}(id) \cdot (Mtx)_U(T) \cdot (Mtx)_{U,W}(id)$$

by the Commutative Diagram Theorem. By Proposition 1, $(Mtx)_{W,U}(id) = ((Mtx)_{U,W}(id))^{-1}$, so $(Mtx)_W(T)$ and $(Mtx)_U(T)$ are similar. For the second bullet, $P$ is invertible, so by Proposition 1, there is a basis $W$ such that $P = (Mtx)_{U,W}(id)$ and $P^{-1} = (Mtx)_{W,U}(id)$. Let $T \in L(X,X)$ be the linear transformation such that $(Mtx)_U(T) = B$. Then

$$\begin{aligned}A &= P^{-1}BP \\ &= (Mtx)_{W,U}(id) \cdot (Mtx)_U(T) \cdot (Mtx)_{U,W}(id) \\ &= (Mtx)_W(id \circ T \circ id) \\ &= (Mtx)_W(T)\end{aligned}$$

by the Commutative Diagram Theorem. ■

## 2 Eigenvalues and Eigenvectors of a Linear Transformation

De la Fuente defines eigenvalue and eigenvector for a matrix. Here, we define an eigenvalue for a linear transformation. For finite-dimensional spaces, we show that $\lambda$ is an eigenvalue of $T$ if and only if $\lambda$ is an eigenvalue for every matrix representation of $T$.

**Definition 3** Let $X$ be a vector space and $T \in L(X, X)$. We say that $\lambda$ is an *eigenvalue* of $T$ and $v \neq 0$ is an *eigenvector* corresponding to $\lambda$ if $T(v) = \lambda v$.

**Theorem 4** *Let $X$ be a finite-dimensional vector space, and $U$ any basis. Then $\lambda$ is an eigenvalue of $T$ if and only if $\lambda$ is an eigenvalue of $(Mtx)_U(T)$. $v$ is an eigenvector of $T$ corresponding to $\lambda$ if and only if $(crd)_U(v)$ is an eigenvector of $(Mtx)_U(T)$ corresponding to $\lambda$.*

**Proof:** By the Commutative Diagram Theorem,

$$\begin{aligned} T(v) = \lambda v &\Leftrightarrow (crd)_U(T(v)) = crd_U(\lambda v) \\ &\Leftrightarrow (Mtx)_U(T)\left((crd)_U(v)\right) = \lambda\left((crd)_U(v)\right) \end{aligned}$$

■

**Theorem 5 (Theorem 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:*

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

**Proof:** Use de la Fuente's Theorem 6.7 and Theorem 4 above.■

# 3 Diagonalization of Symmetric Real Matrices

**Definition 6** Let
$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}$$
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, each basis element has unit length, and distinct basis elements are perpendicular.

**Example 7** The standard basis of $\mathbf{R}^n$ is orthonormal.

**Definition 8** A real $n \times n$ matrix $A$ is *unitary* if $A^\top = A^{-1}$ (here, $A^\top$ denotes the *transpose* of $A$: $\left(A^\top\right)_{ij} = A_{ji}$).

**Theorem 9** *A real $n \times n$ matrix $A$ is unitary if and only if the columns of $A$ are orthonormal.*

**Proof:** Let $\alpha_j$ denote the $j^{th}$ column of $A$.
$$\begin{aligned} A^\top = A^{-1} &\Leftrightarrow A^\top A = I \\ &\Leftrightarrow \alpha_i \cdot \alpha_j = \delta_{ij} \\ &\Leftrightarrow \{\alpha_1, \ldots, \alpha_n\} \text{ is orthonormal} \end{aligned}$$

■

**Remark 10** Let $A$ be unitary and let $V$ be the set of columns of $A$; let $W$ be the standard basis of $\mathbf{R}^n$. Since $A$ is unitary, it is invertible, so $V$ is a basis of $\mathbf{R}^n$ and $A^\top = (Mtx)_{V,W}(id)$, where $id$ is the identity transformation on $\mathbf{R}^n$. Since $V$ is orthonormal, the transformation between bases $W$ and $V$ preserves all geometry; the lengths of vectors, and the angles between vectors, are not changed by changing from one basis to the other.

**Theorem 11** *Let $T \in L(\mathbf{R}^n, \mathbf{R}^n)$, $id$ the identity transformation in $L(\mathbf{R}^n, \mathbf{R}^n)$, and suppose that $(Mtx)_W(T)$, the matrix representation of $T$ with respect to the standard basis $W$, is symmetric. Then there is an orthonormal basis $V = \{v_1, \ldots, v_n\}$ of $\mathbf{R}^n$ consisting of eigenvectors of $T$, so that*
$$(Mtx)_W(T) = (Mtx)_{W,V}(id) \cdot (Mtx)_V(T) \cdot (Mtx)_{V,W}(id)$$

*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. Here is a very brief outline. Let $M = (Mtx)_W(t)$. Since $M$ is a real matrix, its characteristic polynomial $\det((Mtx)_W(T) - \lambda I)$ is an $n^{th}$ degree polynomial with real coefficients. Recall that an $n^{th}$ degree polynomial with coefficients in $\mathbf{R}$ or in $\mathbf{C}$ always has $n$ roots (not necessarily distinct) in $\mathbf{C}$.[1] The inner product in $\mathbf{C}^n$ is defined as follows:

$$x \cdot y = \sum_{j=1}^{n} x_i \cdot \overline{y_j}$$

where $\bar{c}$ denotes the complex conjugate of any $c \in \mathbf{C}$; note that this implies that $x \cdot y = \overline{y \cdot x}$. The usual inner product in $\mathbf{R}^n$ is the restriction of this inner product on $\mathbf{C}^n$ to $\mathbf{R}^n$. Given any complex matrix $A$, define $A^*$ to be the matrix whose $(i,j)^{th}$ entry is $\overline{a_{ji}}$; in other words, $A^*$ is formed by taking the complex conjugate of each element of the transpose of $A$. It is easy to verify that given $x, y \in \mathbf{C}^n$ and a complex $n \times n$ matrix $A$, $Ax \cdot y = x \cdot A^* y$. Since $M$ is real and symmetric, $M^* = M$. If $\lambda \in \mathbf{C}$ is an eigenvalue of $M$, with eigenvector $x \in \mathbf{C}^n$, then

$$\begin{aligned}
\lambda |x|^2 &= \lambda (x \cdot x) \\
&= (\lambda x) \cdot x \\
&= (Mx) \cdot x \\
&= x \cdot (M^* x) \\
&= x \cdot (Mx) \\
&= x \cdot (\lambda x) \\
&= \overline{(\lambda x) \cdot x} \\
&= \overline{\lambda (x \cdot x)} \\
&= \overline{\lambda |x|^2} \\
&= \bar{\lambda} |x|^2
\end{aligned}$$

[1] Although we do not need it here, you should also recall that if the coefficients of the polynomial lie in $\mathbf{R}$, the roots occur in conjugate pairs. This fact is important for the solution of differential equations.

which proves that $\lambda = \bar{\lambda}$, hence $\lambda \in \mathbf{R}$.

Notice that if $M$ is real (not necessarily symmetric) and $\lambda \in \mathbf{R}$ is an eigenvalue, then $\det(M - \lambda I) = 0 \Rightarrow \exists v \in \mathbf{R}^n$ s.t. $(M - \lambda I)v = 0$, so there is at least one real eigenvalue. The fact that $M$ is real and symmetric implies that, if an eigenvalue has multiplicity $m$, one can find $m$ independent real eigenvectors corresponding to that eigenvalue. Thus, there is a basis of $\mathbf{R}^n$ consisting of eigenvectors, hence $M$ is diagonalizable over $\mathbf{R}$. To see that the eigenvectors corresponding to distinct eigenvalues are orthogonal, suppose that $Tx = \lambda x$ and $Ty = \rho y$ with $\rho \neq \lambda$. Then

$$\begin{aligned}
\lambda(x \cdot y) &= (\lambda x) \cdot y \\
&= (Mx) \cdot y \\
&= (Mx)^\top y \\
&= \left(x^\top M^\top\right) y \\
&= \left(x^\top M\right) y \\
&= x^\top (My) \\
&= x^\top (\rho y) \\
&= x \cdot (\rho y) \\
&= \rho(x \cdot y)
\end{aligned}$$

so $(\lambda - \rho)(x \cdot y) = 0$; since $\lambda - \rho \neq 0$, we must have $x \cdot y = 0$.

# 4 Application to Quadratic Forms

In the remainder of this note, we use Theorem 11 to study 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 $A = (\alpha_{ij})$. Then $f(x) = x^\top Ax$. Since $A$ is symmetric, $\mathbf{R}^n$ has an orthonormal basis $V = \{v_1, \ldots, v_n\}$ consisting of eigenvectors of $A$; let

$\lambda_1, \ldots, \lambda_n$ denote the corresponding eigenvalues. Thus, $A = U^\top DU$ where $D$ is diagonal (with diagonal elements $\lambda_1, \ldots, \lambda_n$), and $U = (Mtx)_{V,W}(id)$ is unitary. The columns of $U^\top$ (which, of course, are the rows of $U$) are the coordinates of $v_1, \ldots, v_n$, expressed in terms of the standard basis $W$.

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

This proves the following corollary of Theorem 11.

**Corollary 12** *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 $v_1, \ldots, v_n$.*