MA332 Linear Algebra

Lecture handouts –Part 2

- -Determinant
- -Eigenvalues Eigenvectors
- -Markov Chain
- -Positive Definiteness
- -Constrain Optimisation
- -Linear Programming

MA332 Linear Algebra

Determinants

- Formulas for the determinants
- Properties of Determinant
- Applications of determinants

Determinant➔ a number that associates with every square matrix

Determinant of A ➔ $\det \underline{A}$ or $|A|$

Determinant gives a test for invertibility. .

$$\det\underline{\boldsymbol{A}} = \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

MA332 Linear Algebra

For any square matrix $\underline{\boldsymbol{A}}$, let A_{ij} denote the submatrix formed by deleting the ith row and jth column of $\underline{\boldsymbol{A}}$.

The determinant of a 3x3 matrix is defined by determinants of the 2x2 submatrices A_{ij}.

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}.\det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12}.\det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13}.\det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

$$\Delta = a_{11}.\det A_{11} - a_{12}.\det A_{12} + a_{13}.\det A_{13}$$

The determinant of an ***nxn*** matrix is defined by determinants of ***(n-1)x(n-1)*** submatrices.

MA332 Linear Algebra

A recursive definition of a determinant

For $n \geq 2$, the **determinant** of an $n \times n$ matrix $A = [a_{ij}]$ is given by

$$\det A = a_{11}\det A_{11} - a_{12}\det A_{12} + \cdots + (-1)^{1+n} a_{1n}\det A_{1n}$$

$$= \sum_{j=1}^{n} (-1)^{1+j} a_{1j}\det A_{1j}$$

The determinant of an nxn matrix **A** is the sum of n term of the form $\pm a_{ij}$ det A_{ij} with plus and minus sign alternating where entries a_{11}, a_{12}, ….a_{1n} from the first row of A

MA332 Linear Algebra

EXAMPLE: Compute the determinant of $A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{bmatrix}$

Solution

$$\det A = 1\det\begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix} - 2\det\begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} + 0\det\begin{bmatrix} 3 & -1 \\ 2 & 0 \end{bmatrix}$$

$$= \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ = \_\_\_\_\_$$

Common notation: $\det\begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} = \begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix}$.

So

$$\begin{vmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{vmatrix} = 1\begin{vmatrix} -1 & 2 \\ 0 & 1 \end{vmatrix} - 2\begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} + 0\begin{vmatrix} 3 & -1 \\ 2 & 0 \end{vmatrix}$$

MA332 Linear Algebra

Given $A=\lfloor a_{ij}\rfloor$ The **(i, j)-cofactor** of A is the number C_{ij} where

$$C_{ij} = (-1)^{i+j}\det A_{ij}.$$

$$\det A = a_{11}.c_{11} + a_{12}c_{12} + \ldots\ldots a_{1n}c_{1n}$$

A cofactor expansion across the first row of A

$$\begin{vmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{vmatrix} = 1C_{11} + 2C_{12} + 0C_{13}$$

(cofactor expansion across row 1)

MA332 Linear Algebra

Cofactor expansion theorem

THEOREM 1 The determinant of an $n \times n$ matrix A can be computed by a cofactor expansion across any row or down any column:

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$ (expansion across row i)

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$ (expansion down column j)

Use a matrix of signs to determine $(-1)^{i+j}$

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

MA332 Linear Algebra

EXAMPLE: Compute the determinant of $A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{bmatrix}$ using cofactor expansion down column 3.

Solution

$$\begin{vmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{vmatrix} = 0\begin{vmatrix} 3 & -1 \\ 2 & 0 \end{vmatrix} - 2\begin{vmatrix} 1 & 2 \\ 2 & 0 \end{vmatrix} + 1\begin{vmatrix} 1 & 2 \\ 3 & -1 \end{vmatrix} = 1.$$

Example

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

$\det A=?$

MA332 Linear Algebra

EXAMPLE: Compute the determinant of $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 1 & 5 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 3 & 5 \end{bmatrix}$

Solution

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 1 & 5 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 3 & 5 \end{vmatrix}$$

$$= 1\begin{vmatrix} 2 & 1 & 5 \\ 0 & 2 & 1 \\ 0 & 3 & 5 \end{vmatrix} - 0\begin{vmatrix} 2 & 3 & 4 \\ 0 & 2 & 1 \\ 0 & 3 & 5 \end{vmatrix} + 0\begin{vmatrix} 2 & 3 & 4 \\ 2 & 1 & 5 \\ 0 & 3 & 5 \end{vmatrix} - 0\begin{vmatrix} 2 & 3 & 4 \\ 2 & 1 & 5 \\ 0 & 2 & 1 \end{vmatrix}$$

$$= 1 \cdot 2\begin{vmatrix} 2 & 1 \\ 3 & 5 \end{vmatrix} = 14$$

Triangular Matrices:

$$\begin{bmatrix} * & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ 0 & 0 & \ddots & * & * \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * \end{bmatrix} \qquad \begin{bmatrix} * & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ * & * & \ddots & 0 & 0 \\ * & * & \cdots & * & 0 \\ * & * & \cdots & * & * \end{bmatrix}$$

(upper triangular) (lower triangular)

THEOREM 2: If A is a triangular matrix, then $\det A$ is the product of the main diagonal entries of A.

EXAMPLE:

$$\begin{vmatrix} 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & -3 & 5 \\ 0 & 0 & 0 & 4 \end{vmatrix} = \rule{4cm}{0.4pt} = -24$$

MA332 Linear Algebra

Properties of Determinants

THEOREM 3 Let A be a square matrix.

a. If a multiple of one row of A is added to another row of A to produce a matrix B, then $\det A = \det B$.

b. If two rows of A are interchanged to produce B, then $\det B = -\det A$.

c. If one row of A is multiplied by k to produce B, then $\det B = k \cdot \det A$.

Theorem 3(c) indicates that $\begin{vmatrix} * & * & * \\ -2k & 5k & 4k \\ * & * & * \end{vmatrix} = k \begin{vmatrix} * & * & * \\ -2 & 5 & 4 \\ * & * & * \end{vmatrix}$.

MA332 Linear Algebra

EXAMPLE: Compute $\begin{vmatrix} 2 & 4 & 6 \\ 5 & 6 & 7 \\ 7 & 6 & 10 \end{vmatrix}$

Solution

$$\begin{vmatrix} 2 & 4 & 6 \\ 5 & 6 & 7 \\ 7 & 6 & 10 \end{vmatrix} = 2\begin{vmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 7 & 6 & 10 \end{vmatrix} = 2\begin{vmatrix} 1 & 2 & 3 \\ 0 & -4 & -8 \\ 0 & -8 & -11 \end{vmatrix}$$

$$= 2(-4)\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & -8 & -11 \end{vmatrix} = 2(-4)\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 5 \end{vmatrix}$$

$$= 2(-4)(1)(1)(5) = -40$$

MA332 Linear Algebra

EXAMPLE: Compute $\begin{vmatrix} 2 & 3 & 0 & 1 \\ 4 & 7 & 0 & 3 \\ 7 & 9 & -2 & 4 \\ 1 & 2 & 0 & 4 \end{vmatrix}$ using a combination of row reduction and cofactor expansion.

Solution $\begin{vmatrix} 2 & 3 & 0 & 1 \\ 4 & 7 & 0 & 3 \\ 7 & 9 & -2 & 4 \\ 1 & 2 & 0 & 4 \end{vmatrix} = -2\begin{vmatrix} 2 & 3 & 1 \\ 4 & 7 & 3 \\ 1 & 2 & 4 \end{vmatrix} = -2\begin{vmatrix} 2 & 3 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 4 \end{vmatrix}$

$$= 2\begin{vmatrix} 2 & 3 & 1 \\ 1 & 2 & 4 \\ 0 & 1 & 1 \end{vmatrix} = -2\begin{vmatrix} 1 & 2 & 4 \\ 2 & 3 & 1 \\ 0 & 1 & 1 \end{vmatrix} = -2\begin{vmatrix} 1 & 2 & 4 \\ 0 & -1 & -7 \\ 0 & 1 & 1 \end{vmatrix}$$

$$= -2\begin{vmatrix} 1 & 2 & 4 \\ 0 & -1 & -7 \\ 0 & 0 & -6 \end{vmatrix} = -2(1)(-1)(-6) = -12.$$

EXAMPLE: Compute $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 0 & 0 \\ 2 & 7 & 6 & 10 \\ 2 & 9 & 7 & 11 \end{vmatrix}$.

Solution

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 0 & 0 \\ 2 & 7 & 6 & 10 \\ 2 & 9 & 7 & 11 \end{vmatrix} = 5\begin{vmatrix} 1 & 3 & 4 \\ 2 & 6 & 10 \\ 2 & 7 & 11 \end{vmatrix} = 5\begin{vmatrix} 1 & 3 & 4 \\ 0 & 0 & 2 \\ 2 & 7 & 11 \end{vmatrix}$$

$$= 5\begin{vmatrix} 1 & 3 & 4 \\ 0 & 0 & 2 \\ 0 & 1 & 3 \end{vmatrix} = -5\begin{vmatrix} 1 & 3 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{vmatrix} = \underline{\qquad\qquad} = \underline{\qquad}.$$

Suppose A has been reduced to $U = \begin{bmatrix} \blacksquare & * & * & \cdots & * \\ 0 & \blacksquare & * & \cdots & * \\ 0 & 0 & \blacksquare & \cdots & * \\ 0 & 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \blacksquare \end{bmatrix}$ by row replacements and row interchanges, then

$$\det A = \begin{cases} (-1)^r \begin{pmatrix} \text{product of} \\ \text{pivots in } U \end{pmatrix} & \text{when } A \text{ is invertible} \\ & r \text{ is number of row interchanges} \\ 0 & \text{when } A \text{ is not invertible} \end{cases}$$

THEOREM 4 A square matrix is invertible if and only if $\det A \neq 0$.

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

Add 2 times row 1 to row 3

$$\begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 5 & -3 & -6 \\ -5 & -8 & 0 & 9 \end{bmatrix}$$

Add row 2 to row 3 to obtain

$$\begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 0 & 0 & 0 \\ -5 & -8 & 0 & 9 \end{bmatrix}$$

$$\det A = \det\begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 0 & 0 & 0 \\ -5 & -8 & 0 & 9 \end{bmatrix} = 0$$

THEOREM 5 If A is an $n \times n$ matrix, then $\det A^T = \det A$.

Partial proof (2×2 case)

$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc \quad \text{and}$$

$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix}^T = \det\begin{bmatrix} a & c \\ b & d \end{bmatrix} = ad - bc$$

$$\Rightarrow \det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \det\begin{bmatrix} a & c \\ b & d \end{bmatrix}.$$

(3×3 case)

$$\det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

$$\det\begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix} = a\begin{vmatrix} e & h \\ f & i \end{vmatrix} - b\begin{vmatrix} d & g \\ f & i \end{vmatrix} + c\begin{vmatrix} d & g \\ e & h \end{vmatrix}$$

$$\Rightarrow \det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \det\begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}.$$

Implications of Theorem 5?

Theorem 3 still holds if the word *row* is replaced

with ____________________.

THEOREM 6 (Multiplicative Property)

For $n \times n$ matrices A and B, $\det(AB) = (\det A)(\det B)$.

EXAMPLE: Compute $\det A^3$ if $\det A = 5$.

Solution: $\det A^3 = \det(AAA) = (\det A)(\det A)(\det A)$

$= \_\_\_\_\_\_\_\_\_\_\_\_\_\_ = \_\_\_\_\_\_\_\_.$

EXAMPLE: For $n \times n$ matrices A and B, show that A is singular if $\det B \neq 0$ and $\det AB = 0$.

Solution: *Since*

$$(\det A)(\det B) = \det AB = 0$$

and

$$\det B \neq 0,$$

then $\det A = 0$. Therefore A is singular.

Applications of determinants

Cramer’s Rule

Cramer’s rule can be used to study how the solution of **Ax=b** affected by the changes in the entries of b.

Let A be an invertible $n \times n$ matrix. For any $\mathbf{b}$ in R^n, the unique solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$ has entries given by

$$x_i = \frac{\det A_i(\mathbf{b})}{\det A}, \quad i = 1,2,3\ldots n$$

where

$$A_i(\mathbf{b}) = [\mathbf{a}_1 \;\; . \;\; . \;\; \underset{\substack{\uparrow \\ Col\ i}}{\mathbf{b}} \;\; . \;\; . \;\; \mathbf{a}_n]$$

Example

$$3x_1 - 2x_2 = 6$$
$$-5x_1 + 4x_2 = 8$$

The computation of A^{-1}

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$

The matrix of cofactors on the right side of (4) is called the **adjugate** (or **classical adjoint**) of A, denoted by adj A. (The term *adjoint* also has another meaning in advanced texts on linear transformations.) The next theorem simply restates (4).

The adjoint of A_{nxn} is defined to be the transpose of the matrix of cofactors:

$$\boxed{adjA = \left[C_{ij}(A)\right]^T}$$

$$adjA ==> n \times n$$

Example

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

$$adjA = \begin{bmatrix} 11 & -26 & 46 \\ -8 & 13 & -24 \\ 7 & -13 & 21 \end{bmatrix}^T = \begin{bmatrix} 11 & -8 & 7 \\ -26 & 13 & -13 \\ 46 & -24 & 21 \end{bmatrix}$$

Theorem

An Inverse Formula

Let A be an invertible $n \times n$ matrix. Then

$$A^{-1} = \frac{1}{\det A}\,\text{adj } A$$

Example Find the inverse of the matrix $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix}$.

Solution The nine cofactors are

$$C_{11} = +\begin{vmatrix} -1 & 1 \\ 4 & -2 \end{vmatrix} = -2, \quad C_{12} = -\begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} = 3, \quad C_{13} = +\begin{vmatrix} 1 & -1 \\ 1 & 4 \end{vmatrix} = 5$$

$$C_{21} = -\begin{vmatrix} 1 & 3 \\ 4 & -2 \end{vmatrix} = 14, \quad C_{22} = +\begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix} = -7, \quad C_{23} = -\begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = -7$$

$$C_{31} = +\begin{vmatrix} 1 & 3 \\ -1 & 1 \end{vmatrix} = 4, \quad C_{32} = -\begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = 1, \quad C_{33} = +\begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} = -3$$

The adjugate matrix is the *transpose* of the matrix of cofactors. [For instance, C_{12} gc in the (2, 1) position.] Thus

$$\operatorname{adj} A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}$$

We could compute det A directly, but the following computation provides a check the calculations above *and* produces det A:

$$(\operatorname{adj} A)\cdot A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix} = 14I$$

Since $(\operatorname{adj} A)A = 14I$,

$$A^{-1} = \frac{1}{14}\begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix} = \begin{bmatrix} -1/7 & 1 & 2/7 \\ 3/14 & -1/2 & 1/14 \\ 5/14 & -1/2 & -3/14 \end{bmatrix}$$

Eigenvalues and Eigenvectors

The basic concepts presented here –eigenvectors and Eigenvalues are useful throughout pure and applied mathematics. Eigenvalues are also used to study difference equations and continuous dynamical systems.

EXAMPLE: Let $A = \begin{bmatrix} 0 & -2 \\ -4 & 2 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. **Examine the results of Au and Av**

MA332 Linear Algebra

Solution

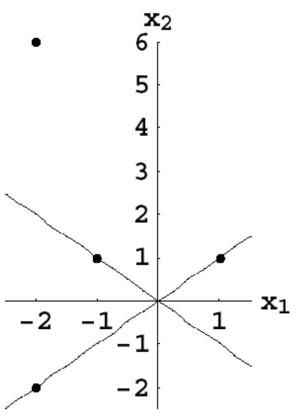
$$A\mathbf{u} = \begin{bmatrix} 0 & -2 \\ -4 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix} = -2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = -2\mathbf{u}$$

$$A\mathbf{v} = \begin{bmatrix} 0 & -2 \\ -4 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 6 \end{bmatrix} \neq \lambda\mathbf{v}$$

$\mathbf{u}$ is called an *eigenvector* of A.

$\mathbf{v}$ is not an eigenvector of A since $A\mathbf{v}$ is not a multiple of $\mathbf{v}$.

MA332 Linear Algebra



$A\mathbf{u} = -2\mathbf{u}$, but $A\mathbf{v} \neq \lambda\mathbf{v}$

MA332 Linear Algebra

DEFINITION

An **eigenvector** of an $n \times n$ matrix A is a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ. A scalar λ is called an **eigenvalue** of A if there is a nontrivial solution $\mathbf{x}$ of $A\mathbf{x} = \lambda\mathbf{x}$; such an $\mathbf{x}$ is called an *eigenvector corresponding to* λ.

EXAMPLE: Show that 4 is an eigenvalue of $A = \begin{bmatrix} 0 & -2 \\ -4 & 2 \end{bmatrix}$ and find the corresponding eigenvectors.

Solution: Scalar 4 is an eigenvalue of A if and only if $A\mathbf{x} = 4\mathbf{x}$ has a nontrivial solution.

$$A\mathbf{x} - 4\mathbf{x} = \mathbf{0}$$

$$A\mathbf{x} - 4(\_\_\_\_)\mathbf{x} = \mathbf{0}$$

$$(A - 4I)\mathbf{x} = \mathbf{0}.$$

To solve $(A-4I)\mathbf{x} = \mathbf{0}$, we need to find $A-4I$ first:

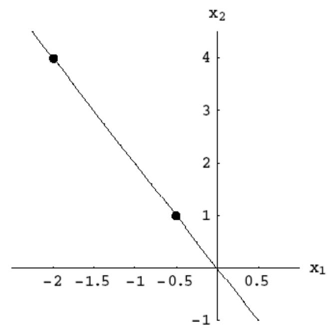
$$A-4I = \begin{bmatrix} 0 & -2 \\ -4 & 2 \end{bmatrix} - \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} -4 & -2 \\ -4 & -2 \end{bmatrix}$$

Now solve $(A-4I)\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} -4 & -2 & 0 \\ -4 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \quad \mathbf{x} = \begin{bmatrix} \frac{-1}{2}x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} \frac{-1}{2} \\ 1 \end{bmatrix}.$$

Each vector of the form $x_2 \begin{bmatrix} \frac{-1}{2} \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda = 4$.



Eigenspace for $\lambda = 4$

Warning: The method just used to find eigen*vectors* ***cannot*** be used to find eigen*values*.

The set of all solutions to $(A-\lambda I)\mathbf{x} = \mathbf{0}$ is called the **eigenspace** of A corresponding to λ.

EXAMPLE: Let $A = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix}$. An eigenvalue of A is $\lambda = 2$. Find a basis for the corresponding eigenspace.

Solution:

$$A-2I = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix} - \begin{bmatrix} \_\_ & 0 & 0 \\ 0 & \_\_ & 0 \\ 0 & 0 & \_\_ \end{bmatrix}$$

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

$$= \begin{bmatrix} \_\_ & 0 & 0 \\ -1 & \_\_ & 1 \\ -1 & 1 & \_\_ \end{bmatrix}$$

Augmented matrix for $(A-2I)\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & 0 \\ -1 & 1 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 + x_3 \\ x_2 \\ x_3 \end{bmatrix} = \_\_\_ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \_\_\_ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

So a basis for the eigenspace corresponding to $\lambda = 2$ is

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

MA332 Linear Algebra

EXAMPLE: Suppose λ is eigenvalue of A. Determine an eigenvalue of A^2 and A^3. In general, what is an eigenvalue of A^n?

Solution: Since λ is eigenvalue of A, there is a nonzero vector **x** such that

$$A\mathbf{x} = \lambda\mathbf{x}.$$

Then

$$\_\_\_A\mathbf{x} = \_\_\_\lambda\mathbf{x}$$

$$A^2\mathbf{x} = \lambda A\mathbf{x}$$

$$A^2\mathbf{x} = \lambda\_\_\_\mathbf{x}$$

$$A^2\mathbf{x} = \lambda^2\mathbf{x}$$

MA332 Linear Algebra

Therefore λ^2 is an eigenvalue of A^2.

Show that λ^3 is an eigenvalue of A^3:

$$\_\_\_A^2\mathbf{x} = \_\_\_\lambda^2\mathbf{x}$$

$$A^3\mathbf{x} = \lambda^2 A\mathbf{x}$$

$$A^3\mathbf{x} = \lambda^3\mathbf{x}$$

Therefore λ^3 is an eigenvalue of A^3.

In general, ______ is an eigenvalue of A^n.

MA332 Linear Algebra

THEOREM 1 The eigenvalues of a triangular matrix are the entries on its main diagonal.

Proof for the 3×3 *Upper Triangular Case*: Let

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}.$$

and then

$$A - \lambda I = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ 0 & a_{22} - \lambda & a_{23} \\ 0 & 0 & a_{33} - \lambda \end{bmatrix}.$$

MA332 Linear Algebra

By definition, λ is an eigenvalue of A if and only if $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution. This occurs if and only if $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has a free variable.

When does this occur?

THEOREM 2 If $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is a linearly independent set.

MA332 Linear Algebra

The characteristic equation

Review

$$A \quad \mathbf{x} \quad = \quad \lambda \quad \mathbf{x}$$

Find eigenvectors $\mathbf{x}$ by solving $(A - \lambda I)\mathbf{x} = \mathbf{0}$.

How do we find the eigenvalues λ?

$\mathbf{x}$ must be nonzero

$\Downarrow$

$(A - \lambda I)\mathbf{x} = \mathbf{0}$ must have nontrivial solutions

$\Downarrow$

$(A - \lambda I)$ is not invertible

$\Downarrow$

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

(called the *characteristic equation*)

Characteristic polynomial: $\det(A - \lambda I)$

Characteristic equation: $\det(A - \lambda I) = 0$

$\boxed{\text{Solve } \det(A - \lambda I) = 0 \text{ for } \lambda \text{ to find the eigenvalues.}}$

MA332 Linear Algebra

EXAMPLE: Find the eigenvalues of $A = \begin{bmatrix} 0 & 1 \\ -6 & 5 \end{bmatrix}$.

Solution: Since

$$A - \lambda I = \begin{bmatrix} 0 & 1 \\ -6 & 5 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} -\lambda & 1 \\ -6 & 5-\lambda \end{bmatrix},$$

the equation $\det(A - \lambda I) = 0$ becomes

$$-\lambda(5-\lambda) + 6 = 0$$

$$\lambda^2 - 5\lambda + 6 = 0$$

Factor:

$$(\lambda - 2)(\lambda - 3) = 0.$$

So the eigenvalues are 2 and 3.

For a 3×3 matrix or larger, recall that a determinant can be computed by cofactor expansion.

MA332 Linear Algebra

EXAMPLE: Find the eigenvalues of $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -5 & 0 \\ 1 & 8 & 1 \end{bmatrix}$.

Solution:

$$A - \lambda I = \begin{bmatrix} 1 - \_\_\_ & 2 & 1 \\ 0 & -5 - \_\_\_ & 0 \\ 1 & 8 & 1 - \_\_\_ \end{bmatrix}$$

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 2 & 1 \\ 0 & -5-\lambda & 0 \\ 1 & 8 & 1-\lambda \end{vmatrix} = (-5-\lambda)\begin{vmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix}$$

$$= (-5-\lambda)\left[(1-\lambda)^2 - 1\right] = (-5-\lambda)[1 - 2\lambda + \lambda^2 - 1]$$

$$= (-5-\lambda)[-2\lambda + \lambda^2] = -(5+\lambda)\lambda[-2+\lambda] = 0$$

$$\Rightarrow \lambda = -5, 0, 2$$

MA332 Linear Algebra

THEOREM

Let A be an $n \times n$ matrix. Then A is invertible if and only if:

- ☐ The number 0 is not an eigenvalue of A.
- ☐ $\det A \neq 0$

Recall that if B is obtained from A by a sequence of row replacements or interchanges, but without scaling, then $\det A = (-1)^r \det B$, where r is the number of row interchanges.

Suppose the echelon form U is obtained from A by a sequence of row replacements or interchanges, but without scaling.

MA332 Linear Algebra

$$A \sim U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & u_{nn} \end{bmatrix}$$

The **determinant** of A, written $\det A$, is defined as follows:

$$\det A = \begin{cases} (-1)^r \cdot \left(\begin{array}{l} \text{product of} \\ \text{pivots in } U \end{array} \right), & \text{when } A \text{ is invertible} \\ 0, & \text{when } A \text{ is not invertible} \end{cases}$$

(r is the number of row interchanges)

MA332 Linear Algebra

EXAMPLE: Find the eigenvalues of $A = \begin{bmatrix} 3 & 2 & 3 \\ 0 & 6 & 10 \\ 0 & 0 & 2 \end{bmatrix}$.

Solution:

$$\det(A - \lambda I) = \det \begin{bmatrix} 3-\lambda & 2 & 3 \\ 0 & 6-\lambda & 10 \\ 0 & 0 & 2-\lambda \end{bmatrix}$$

Characteristic equation:

$$(\qquad)(\qquad)(\qquad) = 0.$$

eigenvalues: _____, _____, _____

The (**algebraic**) **multiplicity** of an eigenvalue is its multiplicity as a root of the characteristic equation.

MA332 Linear Algebra

EXAMPLE: Find the characteristic polynomial of

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 5 & 3 & 0 & 0 \\ 9 & 1 & 3 & 0 \\ 1 & 2 & 5 & -1 \end{bmatrix}$$

and then find all the eigenvalues and the algebraic multiplicity of each eigenvalue.

Solution:

$$\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 0 & 0 & 0 \\ 5 & 3-\lambda & 0 & 0 \\ 9 & 1 & 3-\lambda & 0 \\ 1 & 2 & 5 & -1-\lambda \end{vmatrix}$$

$$= (2-\lambda)(3-\lambda)(3-\lambda)(-1-\lambda) = 0$$

eigenvalues: _____, _____, _____

MA332 Linear Algebra

Similarity

Numerical methods for finding approximating eigenvalues are based upon Theorem 4 to be described shortly.

For $n \times n$ matrices A and B, we say the A is **similar** to B if there is an invertible matrix P such that

$$P^{-1}AP = B \qquad \text{or equivalently,} \qquad A = PBP^{-1}.$$

Theorem 4: If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).

Proof: If $B = P^{-1}AP$, then

$$\det(B - \lambda I) = \det[P^{-1}AP - P^{-1}\lambda IP] = \det[P^{-1}(A - \lambda I)P]$$

$$= \det P^{-1} \cdot \det(A - \lambda I) \cdot \det P = \det(A - \lambda I).$$

MA332 Linear Algebra

Diagonalization

The goal here is to develop a useful factorization $A = PDP^{-1}$, when A is $n \times n$. We can use this to compute A^k quickly for large k.

The matrix D is a *diagonal* matrix (i.e. entries off the main diagonal are all zeros).

D^k is trivial to compute as the following example illustrates.

MA332 Linear Algebra

EXAMPLE: Let $D = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}$. Compute D^2 and D^3. In general, what is D^k, where k is a positive integer?

Solution:

$$D^2 = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} & 0 \\ 0 & \end{bmatrix}$$

$$D^3 = D^2 D = \begin{bmatrix} 5^2 & 0 \\ 0 & 4^2 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} & 0 \\ 0 & \end{bmatrix}$$

MA332 Linear Algebra

$$D^k = \begin{bmatrix} 5^k & 0 \\ 0 & 4^k \end{bmatrix} \quad k \geq 1$$

EXAMPLE: Let $A = \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}$. Find a formula for A^k given that $A = PDP^{-1}$ where $P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, $D = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}$ and $P^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$.

MA332 Linear Algebra

Solution:

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PDDP^{-1} = PD^2P^{-1}$$

Again,

$$A^3 = A^2A = (PD^2P^{-1})(PDP^{-1}) = PD^2(P^{-1}P)DP^{-1} = PD^3P^{-1}$$

In general,

$$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5^k & 0 \\ 0 & 4^k \end{bmatrix}\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \cdot 5^k - 4^k & -5^k + 4^k \\ 2 \cdot 5^k - 2 \cdot 4^k & -5^k + 2 \cdot 4^k \end{bmatrix}.$$

A square matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix, i.e. if $A = PDP^{-1}$ where P is invertible and D is a diagonal matrix.

When is A diagonalizable? (The answer lies in examining the eigenvalues and eigenvectors of A.)

Note that

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Altogether

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 5 & 8 \end{bmatrix}$$

MA332 Linear Algebra

Equivalently,

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} \_\_ & 0 \\ 0 & \_\_ \end{bmatrix}$$

or

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1}$$

In general

$$A\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

MA332 Linear Algebra

and if $\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}$ is invertible, A equals

$$\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}^{-1}$$

THEOREM 5 The Diagonalization Theorem

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

In fact, $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A. In this case, the diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P.

MA332 Linear Algebra

EXAMPLE: Diagonalize the following matrix, if possible.

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

Step 1. Find the eigenvalues of A.

$$\det(A - \lambda I) = \det\begin{bmatrix} 2-\lambda & 0 & 0 \\ 1 & 2-\lambda & 1 \\ -1 & 0 & 1-\lambda \end{bmatrix} = (2-\lambda)^2(1-\lambda) = 0.$$

Eigenvalues of A: $\lambda = 1$ and $\lambda = 2$.

MA332 Linear Algebra

Step 2. Find three linearly independent eigenvectors of A.

By solving $(A - \lambda I)\mathbf{x} = \mathbf{0}$, for each value of λ, we obtain the following:

Basis for $\lambda = 1$: $\mathbf{v}_1 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

,

Basis for $\lambda = 2$: $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ $\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

Step 3: Construct P from the vectors in step 2.

$$P = \begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

MA332 Linear Algebra

Step 4: Construct D from the corresponding eigenvalues.

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Step 5: Check your work by verifying that AP = PD

$$\boldsymbol{AP} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

$$\boldsymbol{PD} = \begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

MA332 Linear Algebra

EXAMPLE: Diagonalize the following matrix, if possible.

$$A = \begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & 2 \\ 0 & 0 & 4 \end{bmatrix}.$$

Since this matrix is triangular, the eigenvalues are $\lambda = 2$ and $\lambda = 4$. By solving $(A - \lambda I)\mathbf{x} = \mathbf{0}$ for each eigenvalue, we would find the following:

Basis for $\lambda = 2$: $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

Basis for $\lambda = 4$: $\mathbf{v}_2 = \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix}$

MA332 Linear Algebra

Every eigenvector of A is a multiple of $\mathbf{v}_1$ or $\mathbf{v}_2$ which means there are not three linearly independent eigenvectors of A and by Theorem 5, A is not diagonalizable.

EXAMPLE: Why is $A = \begin{bmatrix} 2 & 0 & 0 \\ 2 & 6 & 0 \\ 3 & 2 & 1 \end{bmatrix}$ diagonalizable?

Solution: Since A has three eigenvalues ($\lambda_1 =$ ____, $\lambda_2 =$ ____, $\lambda_3 =$ ____) and since eigenvectors corresponding to distinct eigenvalues are linearly independent, A has three linearly independent eigenvectors and it is therefore diagonalizable.

THEOREM 6 An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

MA332 Linear Algebra

EXAMPLE: Diagonalize the following matrix, if possible.

$$A = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 24 & -12 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Solution: Eigenvalues: -2 and 2 (each with multiplicity 2).

Solving $(A-\lambda I)\mathbf{x} = \mathbf{0}$ yields the following eigenspace basis sets.

Basis for $\lambda = -2$: $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -6 \\ 0 \end{bmatrix}$ $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 0 \end{bmatrix}$

MA332 Linear Algebra

Basis for $\lambda = 2$: $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ $\mathbf{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is linearly independent

$\Rightarrow P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 \end{bmatrix}$ is invertible

$\Rightarrow A = PDP^{-1}$, where

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -6 & 3 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

MA332 Linear Algebra

THEOREM 7 Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \ldots, \lambda_p$.

a. For $1 \leq k \leq p$, the dimension of the eigenspace for λ_k is less than or equal to the multiplicity of the eigenvalue λ_k.

b. The matrix A is diagonalizable if and only if the sum of the dimensions of the distinct eigenspaces equals n, and this happens if and only if the dimension of the eigenspace for each λ_k equals the multiplicity of λ_k.

c. If A is diagonalizable and β_k is a basis for the eigenspace corresponding to λ_k for each k, then the total collection of vectors in the sets $\beta_1, \ldots, \beta_p$ forms an eigenvector basis for $\mathbf{R}^n$.

Application to Markov Chains

A Markov model is a mathematical model of a variety of situations in business. It is used to describe an experiment or measurement that perform many times in the same way, where outcome of each trial of the experiment will be one of several specified outcomes and where the outcome of one trial depends on the immediately preceding trial.

e.g. If population of a city and its suburbs were measured each year it can be represented as a vector $x_0 = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$

A probability vector: a vector with nonnegative entries that add up to 1

A stochastic matrix: a square matrix whose column are probability vectors (v)

A markov chain is a sequence of probability vector together with a stochastic matrix .

Markov Matrix

The Markov chain is described by the first order difference equations

$$x_{k+1} = Ax_k \qquad \text{For } k=0,1,2\ldots$$

$\underline{A}$ is a Markov matrix.

Example: Each year 1/10 of the people outside Bangkok move in and 2/10 of the people inside Bangkok move out.

A difference equation

y = numbers of people outside Bangkok

z = numbers of people inside Bangkok

$$y_1 = 0.9y_0 + 0.2z_0$$

$$z_1 = 0.1y_0 + 0.8z_0$$

$$\begin{bmatrix} y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}$$

2 essential properties of a Markov process

The total number stays fixed and numbers of people inside BKK and outside BKK can never become negative.

-Each column of the Markov matrix adds up to 1.

-The Markov matrix has no negative entries.

$$A = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}$$ is a Markov matrix.

$$\det(A - \lambda I) = \lambda^2 - 1.7\lambda + .7 = 0$$

$$\lambda_1 = 1 \quad \text{and} \quad \lambda_2 = 0.7$$

$$A = PDP^{-1} = \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0.7 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}$$

The distribution after k years

$$x_{k+1} = Ax_k$$

$$\begin{bmatrix} y_k \\ z_k \end{bmatrix} = A^k \begin{bmatrix} y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 1^k & \\ & 0.7^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}$$

$$= 1^k (y_0 + z_0) \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} + (y_0 - 2z_0)(0.7)^k \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix}$$

$$= c_1 \lambda_1^k \underline{x}_1 + c_2 \lambda_2^k \underline{x}_2$$

k large $(0.7)^k$ is extremely small . The solution approaches a limiting stage

$$\begin{bmatrix} y_\infty \\ z_\infty \end{bmatrix} = (y_0 + z_0) \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

The total population = $(y_0 + z_0)$

The steady state is the eigenvector of *A* correspoding to λ =1

Summary of the Markov process

A Markov matrix is nonnegative entries, with each column adding to 1and with the following properties:

a) $\lambda_1 = 1$ is an eigenvalue.

b) Its eigenvector x_1 is non-negative –and it is steady state since $A\underline{x}_1 = \underline{x}_1$ $|\lambda_i| \le 1$

c) The other eigenvalues satisfy .

d) If any power of *A* has all positive entries, these other $|\lambda_i| \le 1$ are below 1. The solution $A^k u_0$ approaches a multiple of x_1 which is the steady state u_∞

EXAMPLE 1 In Section 1.10 we examined a model for population movement between a city and its suburbs. See Fig. 1. The annual migration between these two parts of the metropolitan region was governed by the *migration matrix M*:

$$\begin{array}{cc} & \text{From:} \\ & \begin{array}{cc} \text{City} & \text{Suburbs} \end{array} \quad \text{To:} \\ M = & \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{array}{l} \text{City} \\ \text{Suburbs} \end{array} \end{array}$$

That is, each year 5% of the city population moves to the suburbs, and 3% of the suburban population moves to the city. The columns of M are probability vectors, so M is a stochastic matrix. Suppose the 2000 population of the region is 600,000 in the city and 400,000 in the suburbs. Then the initial distribution of the population in the region is given by $\mathbf{x}_0$ in (1) above. What is the distribution of the population in 2001? In 2002?

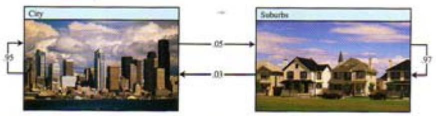


FIGURE 1 Annual percentage migration between city and suburbs.

Solution In Example 3 of Section 1.10, we saw that after one year, the population vector $\begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix}$ changed to

$$\begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix} = \begin{bmatrix} 582{,}000 \\ 418{,}000 \end{bmatrix}$$

If we divide both sides of this equation by the total population of 1 million, and use the fact that $kM\mathbf{x} = M(k\mathbf{x})$, we find that

$$\begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} .600 \\ .400 \end{bmatrix} = \begin{bmatrix} .582 \\ .418 \end{bmatrix}$$

The vector $\mathbf{x}_1 = \begin{bmatrix} .582 \\ .418 \end{bmatrix}$ gives the population distribution in 2001. That is, 58.2% of the region lived in the city and 41.8% lived in the suburbs. Similarly, the population

EXAMPLE 3 Let $P = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}$ and $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. Consider a system whose state is described by the Markov chain $\mathbf{x}_{k+1} = P\mathbf{x}_k$, for $k = 0, 1, \ldots$. What happens to the system as time passes? Compute the state vectors $\mathbf{x}_1, \ldots, \mathbf{x}_{15}$ to find out.

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} .5 \\ .3 \\ .2 \end{bmatrix}$$

$$\mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix} \begin{bmatrix} .5 \\ .3 \\ .2 \end{bmatrix} = \begin{bmatrix} .37 \\ .45 \\ .18 \end{bmatrix}$$

$$\mathbf{x}_3 = P\mathbf{x}_2 = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix} \begin{bmatrix} .37 \\ .45 \\ .18 \end{bmatrix} = \begin{bmatrix} .329 \\ .525 \\ .146 \end{bmatrix}$$

The results of further calculations are shown below, with entries rounded to four or five significant figures.

$$\mathbf{x}_4 = \begin{bmatrix} .3133 \\ .5625 \\ .1242 \end{bmatrix}, \quad \mathbf{x}_5 = \begin{bmatrix} .3064 \\ .5813 \\ .1123 \end{bmatrix}, \quad \mathbf{x}_6 = \begin{bmatrix} .3032 \\ .5906 \\ .1062 \end{bmatrix}, \quad \mathbf{x}_7 = \begin{bmatrix} .3016 \\ .5953 \\ .1031 \end{bmatrix}$$

$$\mathbf{x}_8 = \begin{bmatrix} .3008 \\ .5977 \\ .1016 \end{bmatrix}, \quad \mathbf{x}_9 = \begin{bmatrix} .3004 \\ .5988 \\ .1008 \end{bmatrix}, \quad \mathbf{x}_{10} = \begin{bmatrix} .3002 \\ .5994 \\ .1004 \end{bmatrix}, \quad \mathbf{x}_{11} = \begin{bmatrix} .3001 \\ .5997 \\ .1002 \end{bmatrix}$$

$$\mathbf{x}_{12} = \begin{bmatrix} .30005 \\ .59985 \\ .10010 \end{bmatrix}, \quad \mathbf{x}_{13} = \begin{bmatrix} .30002 \\ .59993 \\ .10005 \end{bmatrix}, \quad \mathbf{x}_{14} = \begin{bmatrix} .30001 \\ .59996 \\ .10002 \end{bmatrix}, \quad \mathbf{x}_{15} = \begin{bmatrix} .30001 \\ .59998 \\ .10001 \end{bmatrix}$$

These vectors seem to be approaching $\mathbf{q} = \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix}$. The probabilities are hardly changing from one value of k to the next. Observe that the following calculation is exact (with no rounding error):

$$P\mathbf{q} = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix} \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix} = \begin{bmatrix} .15 + .12 + .03 \\ .09 + .48 + .03 \\ .06 + 0 + .04 \end{bmatrix} = \begin{bmatrix} .30 \\ .60 \\ .10 \end{bmatrix} = \mathbf{q}$$

When the system is in state $\mathbf{q}$, there is no change in the system from one measurement to the next.

Example

Rent-a-Lemon has three locations from which to rent a car for one day: Airport, downtown and the valley

Daily Migration:

Rented From				
Airport	Downtown	Valley		
.95	.02	.05	Airport	**Returned To**
.03	.90	.05	Downtown	
.02	.08	.90	Valley	

$$M = \begin{bmatrix} .95 & .02 & .05 \\ .03 & .90 & .05 \\ .02 & .08 & .90 \end{bmatrix}$$

(migration matrix)

$$\mathbf{x}_0 = \begin{bmatrix} .5 \\ .3 \\ .2 \end{bmatrix}$$

(initial fraction of cars at *airport*)
(initial fraction of cars *downtown*)
(initial fraction of cars at *valley* location)

(initial distribution vector which is a *probability vector*)

Interpretation of $M\mathbf{x}_0$

$$M\mathbf{x}_0 = \begin{bmatrix} .95 & .02 & .05 \\ .03 & .90 & .05 \\ .02 & .08 & .90 \end{bmatrix}\begin{bmatrix} .5 \\ .3 \\ .2 \end{bmatrix} =$$

$$.5\begin{bmatrix} .95 \\ .03 \\ .02 \end{bmatrix} + .3\begin{bmatrix} .02 \\ .90 \\ .08 \end{bmatrix} + .2\begin{bmatrix} .05 \\ .05 \\ .90 \end{bmatrix}$$

↑	↑	↑
Redistribution of airport cars	Redistribution of downtown cars	Redistribution of valley cars

Distribution after one day=

$$\mathbf{x}_1 = M\mathbf{x}_0 = \begin{bmatrix} .95 & .02 & .05 \\ .03 & .90 & .05 \\ .02 & .08 & .90 \end{bmatrix}\begin{bmatrix} .5 \\ .3 \\ .2 \end{bmatrix} = \begin{bmatrix} 0.491 \\ 0.295 \\ 0.214 \end{bmatrix}$$

$\mathbf{x}_{k+1} = M\mathbf{x}_k$ for $k = 0, 1, 2, \ldots$

(Markov Chain)

Distribution after two days=

$$\mathbf{x}_2 = M\mathbf{x}_1 = \begin{bmatrix} .95 & .02 & .05 \\ .03 & .90 & .05 \\ .02 & .08 & .90 \end{bmatrix}\begin{bmatrix} 0.491 \\ 0.295 \\ 0.214 \end{bmatrix} = \begin{bmatrix} 0.483 \\ 0.290 \\ 0.226 \end{bmatrix}$$

$$\mathbf{x}_3 = M\mathbf{x}_2 = \begin{bmatrix} .95 & .02 & .05 \\ .03 & .90 & .05 \\ .02 & .08 & .90 \end{bmatrix}\begin{bmatrix} 0.483 \\ 0.290 \\ 0.226 \end{bmatrix} = \begin{bmatrix} 0.475 \\ 0.287 \\ 0.236 \end{bmatrix}$$

$$\mathbf{x}_4 = M\mathbf{x}_3 = \begin{bmatrix} .95 & .02 & .05 \\ .03 & .90 & .05 \\ .02 & .08 & .90 \end{bmatrix}\begin{bmatrix} 0.475 \\ 0.287 \\ 0.236 \end{bmatrix} = \begin{bmatrix} 0.468 \\ 0.284 \\ 0.244 \end{bmatrix}$$

$$\vdots$$

$$\mathbf{x}_{49} = M\mathbf{x}_{48} = \begin{bmatrix} 0.417 \\ 0.278 \\ 0.305 \end{bmatrix}$$

$$\mathbf{x}_{50} = M\mathbf{x}_{49} = \begin{bmatrix} 0.417 \\ 0.278 \\ 0.305 \end{bmatrix}$$ (long term distribution)

$$\vdots$$

$$\mathbf{x} = \begin{bmatrix} 0.417 \\ 0.278 \\ 0.305 \end{bmatrix}$$ is called a **steady state vector** since $\mathbf{x} = M\mathbf{x}$

Finding the Steady State Vector

$$M\mathbf{x} = \mathbf{x}$$

$$M\mathbf{x} = I\mathbf{x}$$

$$M\mathbf{x} - I\mathbf{x} = \mathbf{0}$$

$$(M-I)\mathbf{x} = \mathbf{0}$$

Solve $(M-I)\mathbf{x} = \mathbf{0}$ to find the steady state vector. Note that the solution $\mathbf{x}$ must be a *probability vector*.

EXAMPLE: Suppose that 3% of the population of the U.S. lives in the State of Washington. Suppose the migration of the population into and out of Washington State will be constant for many years according to the following migration probabilities. What percentage of the total U.S. population will eventually live in Washington?

From: WA	From: Rest of U.S.	To:
.9	.01	WA
.1	.99	Rest of U.S.

Solution

$$M = \begin{bmatrix} .9 & .01 \\ .1 & .99 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} \text{\% of people in WA} \\ \text{\% in rest of U.S.} \end{bmatrix}$$

$$M - I = \begin{bmatrix} .9 & .01 \\ .1 & .99 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -0.1 & 0.01 \\ 0.1 & -0.01 \end{bmatrix}$$

Solve $(M-I)\mathbf{x} = \mathbf{0}$

$$\begin{bmatrix} -0.1 & 0.01 & 0 \\ 0.1 & -0.01 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -0.1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.1x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 0.1 \\ 1 \end{bmatrix}$$

One solution: $\mathbf{x} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$

Solution we want has entries which add up to one:

$$\mathbf{x} = \begin{bmatrix} 1/11 \\ 10/11 \end{bmatrix} \approx \begin{bmatrix} 0.091 \\ 0.909 \end{bmatrix}$$

EXAMPLE 5 Let $P = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix}$. Find a steady-state vector for P.

Solution First, solve the equation $P\mathbf{x} = \mathbf{x}$.

$$P\mathbf{x} - \mathbf{x} = \mathbf{0}$$
$$P\mathbf{x} - I\mathbf{x} = \mathbf{0} \qquad \text{Recall from Section 1.4 that } I\mathbf{x} = \mathbf{x}.$$
$$(P - I)\mathbf{x} = \mathbf{0}$$

For P as above,

$$P - I = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -.4 & .3 \\ .4 & -.3 \end{bmatrix}$$

To find all solutions of $(P - I)\mathbf{x} = \mathbf{0}$, row reduce the augmented matrix:

$$\begin{bmatrix} -.4 & .3 & 0 \\ .4 & -.3 & 0 \end{bmatrix} \sim \begin{bmatrix} -.4 & .3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -3/4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Then $x_1 = \frac{3}{4}x_2$ and x_2 is free. The general solution is $x_2 \begin{bmatrix} 3/4 \\ 1 \end{bmatrix}$.

Next, choose a simple basis for the solution space. One obvious choice is $\begin{bmatrix} 3/4 \\ 1 \end{bmatrix}$ but a better choice with no fractions is $\mathbf{w} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ (corresponding to $x_2 = 4$).

Finally, find a probability vector in the set of all solutions of $P\mathbf{x} = \mathbf{x}$. This process is easy, since every solution is a multiple of the $\mathbf{w}$ above. Divide $\mathbf{w}$ by the sum of its entries and obtain

$$\mathbf{q} = \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix}$$

As a check, compute

$$P\mathbf{q} = \begin{bmatrix} 6/10 & 3/10 \\ 4/10 & 7/10 \end{bmatrix} \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix} = \begin{bmatrix} 18/70 + 12/70 \\ 12/70 + 28/70 \end{bmatrix} = \begin{bmatrix} 30/70 \\ 40/70 \end{bmatrix} = \mathbf{q}$$

Exercise: Markov chain

1. Multinational companies in the US, Asia, and Europe have assets of \$12 trillion. At the start, \$6 trillion are in the US, \$6 trillion in Europe. Each year half the US money stays home, ¼ each goes to Asia and Europe For Asia and Europe, half stays home and half is sent to the US.

   a) Find the eigenvalues and eigenvectors of this matrix *A*.
   b) What is the limiting distribution of the \$12 trillion as the world ends ($k \rightarrow \alpha$)?

2. A city has only rainy or sunny days. Assume that there is an 80% probability of having a sunny day today, and 80% probability of having sun the day following a sunny day, and a 60% probability of having rain the day following a rainy day. If there is an 80% probability of having a sunny day today, calculate the probabilities of rain and shine for each of the following 10 days and calculate the long term probabilities of rain and shine, and find out how many days it will take to reach this steady state vector with 2 decimals accuracy.

Symmetric Matrices

$$\boxed{A^T = A}$$

- Square matrix
- Main diagonal entries are arbitrary, other entries occur in pairs.

Symmetric: $\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$, $\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & 8 \\ 0 & 8 & -7 \end{bmatrix}$, $\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$

Nonsymmetric: $\begin{bmatrix} 1 & -3 \\ 3 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & -4 & 0 \\ -6 & 1 & -4 \\ 0 & -6 & 1 \end{bmatrix}$, $\begin{bmatrix} 5 & 4 & 3 & 2 \\ 4 & 3 & 2 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}$

What is special when a matrix is symmetry?

EXAMPLE 2 If possible, diagonalize the matrix $A = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$.

Solution The characteristic equation of A is

$$0 = -\lambda^3 + 17\lambda^2 - 90\lambda + 144 = -(\lambda - 8)(\lambda - 6)(\lambda - 3)$$

Standard calculations produce a basis for each eigenspace:

$$\lambda = 8: \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}; \quad \lambda = 6: \mathbf{v}_2 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}; \quad \lambda = 3: \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

These three vectors form a basis for $\mathbb{R}^3$, and we could use them as the columns for a matrix P that diagonalizes Since a nonzero multiple of an eigenvector is still an eigenvector, we can normalize $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ to produce the unit eigenvectors

$$\mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

Let

$$P = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}, \quad D = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Then $A = PDP^{-1}$, as usual. But this time, since P is square and has orthonormal columns, P is an *orthogonal* matrix, and P^{-1} is simply P^T.

THEOREM 1 If A is symmetric, then any two eigenvectors from different eigenspaces are orthogonal.

PROOF Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be eigenvectors that correspond to distinct eigenvalues, say, λ_1 and λ_2. To show that $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, compute

$$\begin{aligned} \lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_2 &= (\lambda_1 \mathbf{v}_1)^T \mathbf{v}_2 = (A\mathbf{v}_1)^T \mathbf{v}_2 && \text{Since } \mathbf{v}_1 \text{ is an eigenvector} \\ &= (\mathbf{v}_1^T A^T)\mathbf{v}_2 = \mathbf{v}_1^T (A\mathbf{v}_2) && \text{Since } A^T = A \\ &= \mathbf{v}_1^T (\lambda_2 \mathbf{v}_2) && \text{Since } \mathbf{v}_2 \text{ is an eigenvector} \\ &= \lambda_2 \mathbf{v}_1^T \mathbf{v}_2 = \lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2 \end{aligned}$$

Hence $(\lambda_1 - \lambda_2)\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. But $\lambda_1 - \lambda_2 \neq 0$, so $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.

The special type of diagonalization in Example 2 is crucial for the theory of symmetric matrices. A matrix A is said to be **orthogonally diagonalizable** if there are an orthogonal matrix P (with $P^{-1} = P^T$) and a diagonal matrix D such that

$$A = PDP^T = PDP^{-1} \tag{1}$$

To orthogonally diagonalize an $n \times n$ matrix, we must find n linearly independent and orthonormal eigenvectors. When is this possible? If A is orthogonally diagonalizable as in (1), then

$$A^T = (PDP^T)^T = P^{TT}D^TP^T = PDP^T = A$$

THEOREM 2 An $n \times n$ matrix A is orthogonally diagonalizable if and only if A is a symmetric matrix.

EXAMPLE 3 Orthogonally diagonalize the matrix $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$, whose characteristic equation is

$$0 = -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda - 7)^2(\lambda + 2)$$

Solution The usual calculations produce bases for the eigenspaces:

$$\lambda = 7: \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = -2: \mathbf{v}_3 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}$$

Although $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent, they are not orthogonal.

the projection of $\mathbf{v}_2$ onto $\mathbf{v}_1$ is $\frac{\mathbf{v}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1$, and the component of $\mathbf{v}_2$ orthogonal to $\mathbf{v}_1$ is

$$\mathbf{z}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} - \frac{-1/2}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix}$$

Normalizing $\mathbf{v}_1$ and $\mathbf{z}_2$, we obtain the following orthonormal basis for the eigenspace for $\lambda = 7$:

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}$$

An orthonormal basis for the eigenspace for $\lambda = -2$ is

$$\mathbf{u}_3 = \frac{1}{\|2\mathbf{v}_3\|}2\mathbf{v}_3 = \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$$

By Theorem 1, $\mathbf{u}_3$ is orthogonal to the other eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$. Hence $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal set. Let

$$P = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3\,] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \end{bmatrix}, \quad D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

THEOREM 3 The Spectral Theorem for Symmetric Matrices

An $n \times n$ symmetric matrix A has the following properties:

a. A has n real eigenvalues, counting multiplicities.
b. The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation.
c. The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal.
d. A is orthogonally diagonalizable.

Spectral Decomposition

Suppose $A = PDP^{-1}$, where the columns of P are orthonormal eigenvectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ of A and the corresponding eigenvalues $\lambda_1, \dots, \lambda_n$ are in the diagonal matrix D. Then, since $P^{-1} = P^T$,

$$A = PDP^T = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_n] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

$$= [\lambda_1\mathbf{u}_1 \ \cdots \ \lambda_n\mathbf{u}_n] \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

Using the column–row expansion of a product

$$A = \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \lambda_2\mathbf{u}_2\mathbf{u}_2^T + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T \qquad (2)$$

EXAMPLE 4 Construct a spectral decomposition of the matrix A that has the orthogonal diagonalization

$$A = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 8 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

Solution Denote the columns of P by $\mathbf{u}_1$ and $\mathbf{u}_2$. Then

$$A = 8\mathbf{u}_1\mathbf{u}_1^T + 3\mathbf{u}_2\mathbf{u}_2^T$$

To verify this decomposition of A, compute

$$\mathbf{u}_1\mathbf{u}_1^T = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} [2/\sqrt{5} \quad 1/\sqrt{5}] = \begin{bmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{bmatrix}$$

$$\mathbf{u}_2\mathbf{u}_2^T = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} [-1/\sqrt{5} \quad 2/\sqrt{5}] = \begin{bmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{bmatrix}$$

and

$$8\mathbf{u}_1\mathbf{u}_1^T + 3\mathbf{u}_2\mathbf{u}_2^T = \begin{bmatrix} 32/5 & 16/5 \\ 16/5 & 8/5 \end{bmatrix} + \begin{bmatrix} 3/5 & -6/5 \\ -6/5 & 12/5 \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = A$$

Quadratic forms

A **quadratic form** on $\mathbb{R}^n$ is a function Q defined on $\mathbb{R}^n$ whose value at a vector $\mathbf{x}$ in $\mathbb{R}^n$ can be computed by an expression of the form $Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$, where A is an $n \times n$ symmetric matrix. The matrix A is called the **matrix of the quadratic form.**

e.g. a nonzero quadratic form $Q(\mathbf{x}) = \mathbf{x}^TI\mathbf{x} = \|\mathbf{x}\|^2$.

The connection between symmetric matrix A and the quadratic form $\mathbf{X}^T\mathbf{AX}$

EXAMPLE 1 Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Compute $\mathbf{x}^TA\mathbf{x}$ for the following matrices:

a. $A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$ b. $A = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$

Solution

a. $\mathbf{x}^TA\mathbf{x} = [x_1 \quad x_2] \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [x_1 \quad x_2] \begin{bmatrix} 4x_1 \\ 3x_2 \end{bmatrix} = 4x_1^2 + 3x_2^2.$

b. There are two -2 entries in A. Watch how they enter the calculations. The (1, 2)-entry in A is in boldface type.

$$\mathbf{x}^TA\mathbf{x} = [x_1 \quad x_2] \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [x_1 \quad x_2] \begin{bmatrix} 3x_1 - 2x_2 \\ -2x_1 + 7x_2 \end{bmatrix}$$

$$= x_1(3x_1 - 2x_2) + x_2(-2x_1 + 7x_2)$$

$$= 3x_1^2 - 2x_1x_2 - 2x_2x_1 + 7x_2^2$$

$$= 3x_1^2 - 4x_1x_2 + 7x_2^2$$

The presence of $-4x_1x_2$ in the quadratic form in Example 1(b) is due to the -2 entries off the diagonal in the matrix A. In contrast, the quadratic form associated with the diagonal matrix A in Example 1(a) has no x_1x_2 *cross-product* term.

EXAMPLE 2 For $\mathbf{x}$ in $\mathbb{R}^3$, let $Q(\mathbf{x}) = 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3$. Write this quadratic form as $\mathbf{x}^TA\mathbf{x}$.

Solution The coefficients of x_1^2, x_2^2, x_3^2 go on the diagonal of A. To make A symmetric, the coefficient of x_ix_j for $i \neq j$ must be split evenly between the (i, j)- and (j, i)-entries in A. The coefficient of x_1x_3 is 0. It is readily checked that

$$Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x} = [x_1 \quad x_2 \quad x_3] \begin{bmatrix} 5 & -1/2 & 0 \\ -1/2 & 3 & 4 \\ 0 & 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

EXAMPLE 3 Let $Q(\mathbf{x}) = x_1^2 - 8x_1x_2 - 5x_2^2$. Compute the value of $Q(\mathbf{x})$ for $\mathbf{x} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$.

Solution

$$Q(-3, 1) = (-3)^2 - 8(-3)(1) - 5(1)^2 = 28$$
$$Q(2, -2) = (2)^2 - 8(2)(-2) - 5(-2)^2 = 16$$
$$Q(1, -3) = (1)^2 - 8(1)(-3) - 5(-3)^2 = -20$$

Positive definite matrices

Symmetric matrices always have real eigenvalues

The signs of the eigenvalues are important.

Are these eigenvalues positive or negative?

$$\det(A - \lambda I) = 0$$ ok for 2x2 or 3x3

Require ➡ A test to guarantee that eigenvalues of a symmetric matrix are all positive without computing its eigenvalues. ➡ 3 most basic ideas

- Pivots
- Determinants
- Eigenvalues

Minima, maxima and saddle points

Example: Optimization problem → Identifying a minimum

Recap. calculus:

1 variable $f(x)$

$\frac{\partial f}{\partial x} = 0$ ➡ A stationary point

$\frac{\partial^2 f}{\partial x^2} > 0$ ➡ A minimum

2 variables $f(x, y)$

$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$ ➡ A stationary point

$\begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{vmatrix} > 0$ and $\frac{\partial^2 f}{\partial x^2} > 0$ A minimum

Every quadratic form $f(x, y) = ax^2 + 2bxy + cy^2$ has a stationary point a the origin,where $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$. If it has a local minimum at x=y=0, then that point is also a global minimum.

$f''(x, y) = 0$ ➡ The 2nd derivative fails to give a definite decision.

For a true minimum ,*f* is allowed to vanish only at x=y=0. When *f* is strictly positive at all other points it is called **positive definite**.

What are the conditions on *a,b,c* which ensure that $f(x, y) = ax^2 + 2bxy + cy^2$ is positive definite?

Completing the square→ $f(x, y) = ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \left(c - \frac{b^2}{a}\right)y^2$

The quadratic form $f(x, y) = ax^2 + 2bxy + cy^2$ is positive definite if and only if $a > 0$ and $ac - b^2 > 0$. Correspondingly, F has a (nonsigular) minimum at $x = y = 0$ If and only if its first derivatives are zero and

$$\frac{\partial^2 F}{\partial x^2}(0,0) > 0 \text{ and } \left[\frac{\partial^2 F}{\partial x^2}(0,0)\right]\left[\frac{\partial^2 F}{\partial y^2}(0,0)\right] > \left[\frac{\partial^2 F}{\partial x \partial y}(0,0)\right]^2$$

The condition for a maximum (f is maximum whenever –f has a minimum):
The quadratic form is **negative definite** if and only if $a<0$ and $ac-b^2>0$.
The same change applies to F.

If $ac-b^2=0$ ➡ positive semidefinite if $a>0$
➡ negative semidefinite if $a<0$

$f''(x,y)=0$ the combination $ac-b^2<0$

When the size of b dominate a and c, the x and y directions give opposite results.

The quadratic forms are said to be **indefinite**, because they can take either sign; both *f>0 or f<0* are possible depending on x and y.

A stationary point is neither minimum or maximum ➔ a saddle point

A quardratic form Q is:

a) Positive definite if $Q(x_1..x_n)>0$ for all $x_i \neq 0$
b) Negative definite if $Q(x_1..x_n)<0$ for all $x_i \neq 0$
c) Indefinite if $Q(x_1..x_n)$ assumes both positive and negative values

Write a quadratic form in term of the matrix product:

Symmetric matrix

$$f(x,y)=ax^2+2bxy+cy^2=\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

$$\boxed{Q=x^T Ax}$$

(generalise to n dimensions)

For any symmetric matrix *A*, the product $Q=x^T Ax$ is a pure quadratic form:

$$x^T Ax=\begin{bmatrix} x_1 & . & . & x_n \end{bmatrix}\begin{bmatrix} a_{11} & . & . & a_{1n} \\ . & . & . & . \\ . & . & . & . \\ a_{n1} & . & . & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ . \\ . \\ x_n \end{bmatrix}$$

$$x^T Ax=a_{11}x_1^2+a_{12}x_1x_2+a_{21}x_2x_1+\ldots+a_{nn}x_n^2$$
$$=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j.$$

Tests for positive definiteness

A (pure) quadratic form $Q=x^T Ax$ is positive definite if $x^T Ax>0$

Which symmetric matrices have the property $x^T Ax>0$ for all nonzero vectors x?

$$A=\begin{bmatrix} a & b \\ b & c \end{bmatrix}$$; we need $a>0$ and $ac-b^2>0$.

- *Generalize these conditions to a matrix of order n*
- *Find connection with the signs of the eigenvalues*

2 by 2 cases ➔ both eigenvalues are positive

$$ac-b^2 = \det A=\lambda_1\lambda_2$$
$$a+c=\text{trace of } A=\lambda_1+\lambda_2$$

$$Q = x^T A x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = ax_1^2 + 2bx_1x_2 + cx_2^2$$

Complete test for positive definite

i) $x^T A x > 0$

ii) $\lambda_1 > 0$ and $\lambda_2 > 0$

We have 2 conditions for A to be positive $a > 0$ and $ac - b^2 > 0$.

Find connection of these conditions with the sign of eigenvalues.

since $ac - b^2$ is the determinant.

$$\lambda_1 \lambda_2 = ac - b^2 > 0$$

$$\therefore \lambda_1 \lambda_2 > 0 \quad (\lambda_1 \text{ and } \lambda_2 \text{ have the same sign})$$

Trace of A $\lambda_1 + \lambda_2 = a + c > 0$

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

All eigenvalues are positive

iii) $a > 0$ and $ac - b^2 > 0$.

$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ $\det A > 0$ and $\det(A_{11}) > 0$.

All sub-determinants are positive.

iv) $a > 0$ and $\dfrac{ac - b^2}{a} > 0$.

since $ax_1^2 + 2bx_1x_2 + cx_2^2 = a\left(x_1 + \dfrac{b}{a}x_2\right)^2 + \left(\dfrac{ac - b^2}{a}\right)x_2^2$

$$\therefore a > 0 \text{ and } \frac{ac - b^2}{a} > 0$$

both a and $\dfrac{ac - b^2}{a}$ are pivots of A

All pivots are positive.

Each of the following tests is a necessary sufficient condition for the real symmetric matrix A to be positive definite:

i) $x^T A x > 0$ for all nonzero vector x.

ii) All eigenvalues of A satisfy $\lambda_i > 0$.

iii) All the upper left submatrices A_k have positive determinants.

iv) All the pivots (without row exchanges) satisfy $d_i > 0$.

Example Is $A = \begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix}$ a positive definite matrix?

Determinant test
Eigenvalue test
Pivots test
$x^T A x > 0$

Example Is $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ a positive definite matrix?

What is the quadratic function associate with it?

All subdeterminants ➔ 2, 3 and 4

Pivots ➔ 2, 3/2, 4/3

Eigenvalues ➔ $2 - \sqrt{2}$, 2, $2 + \sqrt{2}$

$x^T A x > 0$

$$Q(x) = x^T A x = 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_2x_3$$

Example Is $Q(x) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$ positive definite ?

Constrained optimization

Optimization problems➔ find the minimum or maximum of a quadratic form Q(x) for vector x in some specified set.

Usually problem can be arranged so that x varies over the set of unit vectors.

Example Find the maximum and minimum values of $Q(x) = 9x_1^2 + 4x_2^2 + 3x_3^2$ subjected to the constraint $x^T x = 1$

Since x_2^2 and x_3^2 are nonnegative, note that $4x_2^2 \le 9x_2^2$ and $3x_3^2 \le 9x_3^2$

Maximum

$$Q(x) = 9x_1^2 + 4x_2^2 + 3x_3^2 \le 9x_1^2 + 9x_2^2 + 9x_3^2$$
$$= 9(x_1^2 + x_2^2 + x_3^2)$$
$$= 9 \Leftarrow \quad \textit{Maximum value of } Q(x) \textit{ for } \quad x^T x = 1$$

Minimum

$$Q(x) = 9x_1^2 + 4x_2^2 + 3x_3^2 \ge 3x_1^2 + 3x_2^2 + 3x_3^2$$
$$= 3(x_1^2 + x_2^2 + x_3^2)$$
$$= 3 \Leftarrow \quad \textit{Minimum value of } Q(x) \textit{ for } \quad x^T x = 1$$

$$Q(x) = 9x_1^2 + 4x_2^2 + 3x_3^2$$

$$Q(x) = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Eigenvalues = 9, 4, 3

The greatest and least eigenvalues = maximum and minimum of Q(x)

Example 2

Let $A = \begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$, and let $Q(x) = x^T Ax$ for x in R^2

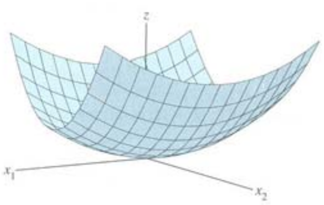


FIGURE 1 $z = 3x_1^2 + 7x_2^2$.

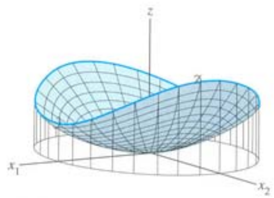


FIGURE 2 The intersection of $z = 3x_1^2 + 7x_2^2$ and the cylinder $x_1^2 + x_2^2 = 1$.

Theorem 4

Let A be a symmetric matrix, and define m and M as in (2). Then M is the greatest eigenvalue λ_1 of A and m is the least eigenvalue of A. The value of $\mathbf{x}^T A\mathbf{x}$ is M when $\mathbf{x}$ is a unit eigenvector $\mathbf{u}_1$ corresponding to M. The value of $\mathbf{x}^T A\mathbf{x}$ is m when $\mathbf{x}$ is a unit eigenvector corresponding to m.

PROOF Orthogonally diagonalize A as PDP^{-1}. We know that

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} \quad \text{when } \mathbf{x} = P\mathbf{y} \tag{3}$$

Also,

$$\|\mathbf{x}\| = \|P\mathbf{y}\| = \|\mathbf{y}\| \quad \text{for all } \mathbf{y}$$

because $P^TP = I$ and $\|P\mathbf{y}\|^2 = (P\mathbf{y})^T(P\mathbf{y}) = \mathbf{y}^TP^TP\mathbf{y} = \mathbf{y}^T\mathbf{y} = \|\mathbf{y}\|^2$. In particular, $\|\mathbf{y}\| = 1$ if and only if $\|\mathbf{x}\| = 1$. Thus $\mathbf{x}^T A\mathbf{x}$ and $\mathbf{y}^T D\mathbf{y}$ assume the same set of values as $\mathbf{x}$ and $\mathbf{y}$ range over the set of all unit vectors.

To simplify notation, we will suppose that A is a 3×3 matrix with eigenvalues $a \ge b \ge c$. Arrange the (eigenvector) columns of P so that $P = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3\,]$ and

$$D = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

Given any unit vector $\mathbf{y}$ in $\mathbb{R}^3$ with coordinates y_1, y_2, y_3, observe that

$$\begin{aligned} ay_1^2 &= ay_1^2 \\ by_2^2 &\le ay_2^2 \\ cy_3^2 &\le ay_3^2 \end{aligned}$$

Adding these inequalities, we have

$$\begin{aligned} \mathbf{y}^T D\mathbf{y} &= ay_1^2 + by_2^2 + cy_3^2 \\ &\le ay_1^2 + ay_2^2 + ay_3^2 \\ &= a(y_1^2 + y_2^2 + y_3^2) \\ &= a\|\mathbf{y}\|^2 = a \end{aligned}$$

Thus $M \le a$, by definition of M. However, $\mathbf{y}^T D\mathbf{y} = a$ when $\mathbf{y} = \mathbf{e}_1 = (1, 0, 0)$, so in fact $M = a$. By (3), the $\mathbf{x}$ that corresponds to $\mathbf{y} = \mathbf{e}_1$ is the eigenvector $\mathbf{u}_1$ of A, because

$$\mathbf{x} = P\mathbf{e}_1 = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3\,]\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{u}_1$$

Thus $M = a = \mathbf{e}_1^T D\mathbf{e}_1 = \mathbf{u}_1^T A\mathbf{u}_1$, which proves the statement about M. A similar argument shows that m is the least eigenvalue, c, and this value of $\mathbf{x}^T A\mathbf{x}$ is attained when $\mathbf{x} = P\mathbf{e}_3 = \mathbf{u}_3$.

Theorem 5

Let A, λ_1, and $\mathbf{u}_1$ be as in Theorem 4 Then the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0$$

is the second greatest eigenvalue, λ_2, and this maximum is attained when $\mathbf{x}$ is an eigenvector $\mathbf{u}_2$ corresponding to λ_2.

EXAMPLE 4 Find the maximum value of $9x_1^2 + 4x_2^2 + 3x_3^2$ subject to the constraints $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^T\mathbf{u}_1 = 0$, where $\mathbf{u}_1 = (1, 0, 0)$. Note that $\mathbf{u}_1$ is a unit eigenvector corresponding to the greatest eigenvalue $\lambda = 9$ of the matrix of the quadratic form.

Solution If the coordinates of $\mathbf{x}$ are x_1, x_2, x_3, then the constraint $\mathbf{x}^T\mathbf{u}_1 = 0$ means simply that $x_1 = 0$. For such a unit vector, $x_2^2 + x_3^2 = 1$, and

$$\begin{aligned} 9x_1^2 + 4x_2^2 + 3x_3^2 &= 4x_2^2 + 3x_3^2 \\ &\le 4x_2^2 + 4x_3^2 \\ &= 4(x_2^2 + x_3^2) \\ &= 4 \end{aligned}$$

Thus the constrained maximum of the quadratic form does not exceed 4. And this value is attained for $\mathbf{x} = (0, 1, 0)$, which is an eigenvector for the second greatest eigenvalue of the matrix of the quadratic form.

EXAMPLE 5 Let A be the matrix in Example 3 and let $\mathbf{u}_1$ be a unit eigenvector corresponding to the greatest eigenvalue of A. Find the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the conditions

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0 \tag{4}$$

Solution From Example 3, the second greatest eigenvalue of A is $\lambda = 3$. Solve $(A - 3I)\mathbf{x} = \mathbf{0}$ to find an eigenvector, and normalize it to obtain

$$\mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}$$

The vector $\mathbf{u}_2$ is automatically orthogonal to $\mathbf{u}_1$ because the vectors correspond to different eigenvalues. Thus the maximum of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints in (4) is 3, attained when $\mathbf{x} = \mathbf{u}_2$.

Theorem 6

Let A be a symmetric $n \times n$ matrix with an orthogonal diagonalization $A = PDP^{-1}$, where the entries on the diagonal of D are arranged so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and where the columns of P are corresponding unit eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Then for $k = 2, \ldots, n$, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0, \quad \ldots, \quad \mathbf{x}^T\mathbf{u}_{k-1} = 0$$

is the eigenvalue λ_k, and this maximum is attained at $\mathbf{x} = \mathbf{u}_k$.

EXAMPLE 6 During the next year, a county government is planning to repair x hundred miles of public roads and bridges and to improve y hundred acres of parks and recreation areas. The county must decide how to allocate its resources (funds, equipment, labor, etc.) between these two projects. If it is more cost-effective to work simultaneously on both projects rather than on only one, then x and y might satisfy a *constraint* such as

$$4x^2 + 9y^2 \leq 36$$

See Fig. 3. Each point (x, y) in the shaded *feasible set* represents a possible public works schedule for the year. The points on the constraint curve, $4x^2 + 9y^2 = 36$, use the maximum amounts of resources available.

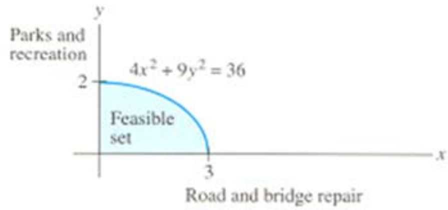


FIGURE 3 Public works schedules.

In choosing its public works schedule, the county wants to consider the opinions of the county residents. To measure the value, or *utility*, that the residents would assign to the various work schedules (x, y), economists sometimes use a function such as

$$q(x, y) = xy$$

The set of points (x, y) at which $q(x, y)$ is a constant is called an *indifference curve*. Three such curves are shown in Fig. 4. Points along an indifference curve correspond to alternatives that county residents as a group would find equally valuable.[2] Find the public works schedule that maximizes the utility function q.

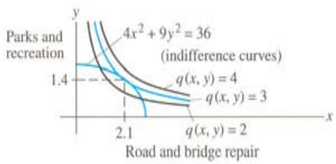


FIGURE 4 The optimum public works schedule is (2.1, 1.4).

Linear Programming (LP)

- An optimization approach
- Constrained optimization
- Cost function and the constraints are linear
- Dealing with rectangular matrice
- The algebra of linear inequalities

Linear inequalities

An inequality divides n dimensional spaces into 2 halfspaces

e.g $x+2y \geq 4$ Boundary is a line in a 2D plane

$x+2y+z \geq 4$ Boundary is a plane in a 3D space

$c_1x_1+......c_nx_n \geq 4$ Boundary is an (n-1) dimensional plane in an nD space

Another constraint which is fundamental in LP is nonnegative x and y

$x \geq 0$ and $y \geq 0$

Example

Suppose General motors makes a profit of $100 on each Chevrolet, $200 on each Buick, and $400 on each Cadillac. They get 20,17 and 14 miles per gallon, respectively, and Congress insists that the average car produced must get 18. The plant can assemble a Chevrolet in 1 minute, a Buick in 2 minutes, and a Cadillac in 3 minutes. What is the maximum profit in an 8-hour day?

Problem:
Maximize $100x+200y+400z$

subject to

$$20x+17y+14z \geq 18(x+y+z)$$
$$x+2y+3z \leq 480$$
$$x, y, z \geq 0$$

Example of Gas production

Resource	Product: Regular (x_1)	Product: Premium (x_2)	Resource available
Raw gas	7m³/ tonne	11m³/ tonne	77m³/ week
Production time	11 hr/tonne	8 hr/tonne	88 hr/week
Storage	9 tonnes	6 tonnes	
Profit	150/ tonne	175/ tonne	

How much of each gas to produce to maximize profits?

Total profit $z = 150x_1 + 175x_2$

Subject to $7x_1 + 11x_2 \leq 77$ (Material constraint)

$11x_1 + 8x_2 \leq 80$ (time constraint)

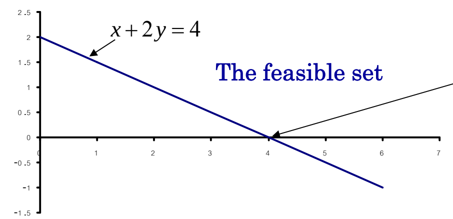
$x_1 \leq 9$ (regular storage constraint)

$x_2 \leq 6$ (premium storage constraint)

$x_1, x_2 \geq 0$ (positivity constraints)

The feasible set and the cost function

(or the objective function)



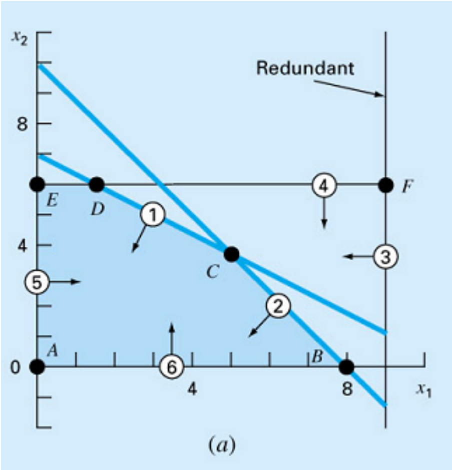
A feasible set is composed of the solutions to a family of linear inequalities

m inequalities

$Ax \geq b$ - intersection of m halfspaces

If requires $x \geq 0$ - add n more halfspaces

Feasible set for the gas production problem



$7x_1 + 11x_2 \le 77$ (Material constraint)

$11x_1 + 8x_2 \le 80$ (time constraint)

$x_1 \le 9$ (regular storage constraint)

$x_2 \le 6$ (premium storage constraint)

$x_1, x_2 \ge 0$ (positivity constraints)

In optimization problem ➔ the particular point that maximizes or minimize a certain cost function (the objective function)

A cost function➔ a function that we would like to maxzimise or minimise its value
e.g net profit (maximize)
total cost of production (minimize)

Example

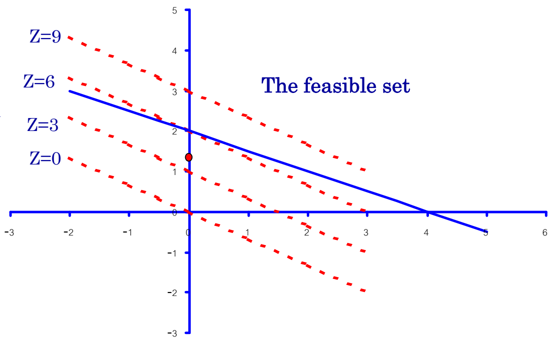
A cost function: $Z = 2x + 3y$ (a family of parallel lines)

Constraints

$$x + 2y \ge 4$$
$$x \ge 0$$
$$y \ge 0$$

Find point (x,y) that lies in the feasible set and minimize the cost function.
➔The first line of cost function (from costs=0) that intersect the feasible set.

Cost function

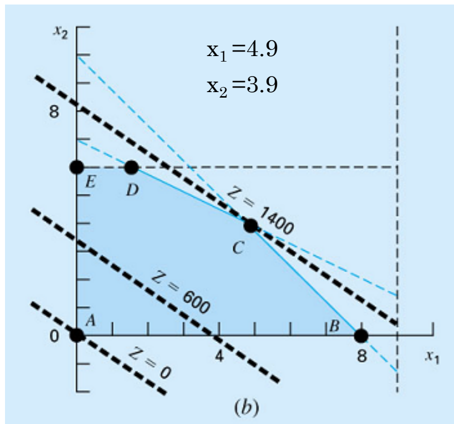


The vector (0,2) is called feasible➔ lies in the feasible set and minimizes the cost function
The optimal vector occurs at a corner of the feasible set

With different cost functions the intersection may not be just a single point
if the cost function Z=x+2y ➔ an infinity of optimal vectors along the edge

Graphical method

Z_{max}=1400



Total profit $z = 150x_1 + 175x_2$

Subject to $7x_1 + 11x_2 \le 77$ (Material constraint)

$11x_1 + 8x_2 \le 80$ (time constraint)

$x_1 \le 9$ (regular storage constraint)

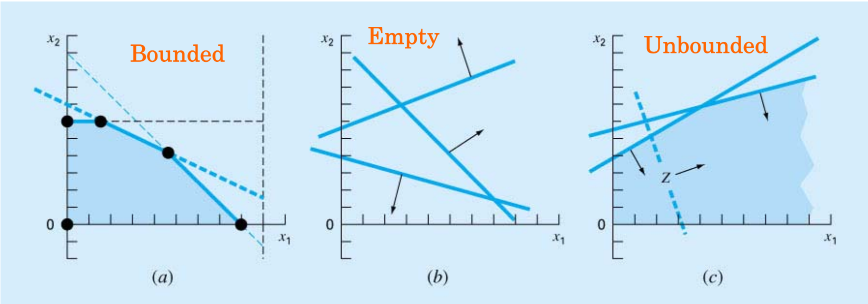
$x_2 \le 6$ (premium storage constraint)

$x_1, x_2 \ge 0$ (positivity constraints)

Graphical method
- Plot constraints
- Plot cost functions

Possible outcomes

Every linear programming problems falls into one of 3 possible categories:



The cost function has **a minimum or a maximum** on the feasible set.

The feasible set is empty
Unsolvable problem
"overconstrainted"
No solution can satisfy all the constrains

The cost function is unbounded on the feasible set.
"underconstrainted"

Slack variables

Introduce slack variables to change inequalities to equality equations

$$x + 2y \geq 4$$
$$w = x + 2y - 4$$
$$w \geq 0$$

The simplex method

- A systematic way to solve LP problems
- The optimal solution ➔ at the corner (the extreme point)
- Reformulate constraint equations as equalities by introducing slack variables.
- n unknowns m constraints

Example 1

A cost function: $Z = 2x + 3y$

Constraints

$x + 2y \geq 4$ — A single constraint

$x \geq 0$
$y \geq 0$ — 2 nonnegative variables

n=2 & m=1

Put the problem directly in matrix form, we are given

1. An m by n matrix A
2. A column vector b with m components
3. A row vector c with n components

To be feasible a vector x has to satisfy $Ax \geq b$ and $x \geq 0$

The optimal vector is the feasible vector with the least cost.

Minimal problem: Minimize cx subject to $Ax \geq b$ and $x \geq 0$

The feasible set is the intersection of m+n halfspaces

Idea of the simplex method

- Locates a corner of the feasible set
- Move from corner to corner along the edge of the feasible set
- Stop when the optimal corner (with min/max cost) is reached

Example 2

$Z = x + y$

constraints

$2x + y \geq 6$

$x + 2y \geq 6$

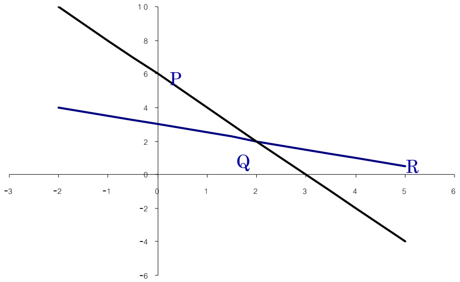
m=2

Nonegativity

$x \geq 0$

$y \geq 0$

n=2



6 possible intersections , 3 are actual corners of the feasible set.

Rewrite $Ax \geq b \longrightarrow w = Ax - b$

The constraints $Ax \geq b$ are translated into $w_1 \geq 0,\ w_2 \geq 0.....,w_m \geq 0$ with one slack variable for every row of A

Example

Mininise: $Z = x + y$

$2x + y \geq 6$

$x + 2y \geq 6$

$x \geq 0, y \geq 0$

The Tableua

Basic	Z	x	y	w	v	solution	Intercept
Z	1	-1	-1	0	0	0	
w	0	2	1	-1	0	6	3
v	0	1	2	0	-1	6	6

Start at x=0, y=0

Basic variables = w, v

Free variables = x , y

- Choose Entering variable = x

  Leaving variable = w

- Replace w with x

Basic	Z	x	y	w	v	solution	Intercept
Z	1	-1	-1	0	0	0	
x	0	1	½	-1/2	0	3	6
v	0	1	2	0	-1	6	

- Perform row operation (Guass-Jordan Method)

Basic	Z	x	y	w	v	solution	Intercept
Z	1	0	-1/2	-1/2	0	3	
x	0	1	-1/2	-1/2	0	3	6
v	0	0	3/2	½	-1	3	2

2^{nd} corner ➔ $x=3, y=0, z=3$

- Choose Entering variable = y

  Leaving variable = v
- Replace v with y
- Perform row operation

Basic	Z	x	y	w	v	solution	Intercept
Z	1	0	0	-1/3	-1/3	4	
x	0	1	0	-2/3	1/3	2	
y	0	0	1	1/3	-2/3	2	

3^{nd} corner ➔ $x=2, y=2, z=4$ (a mimimum point)

Optimal corner is at $x=2, y=2$
The minimum is $Z=4$

Alternatively we can start at a different corner point.

Basic	Z	x	y	w	v	solution	Intercept
Z	1	-1	-1	0	0	0	
w	0	2	1	-1	0	6	3
y	0	1	2	0	-1	6	6

Start at $x=0, y=b$

Basic variables = y, w

Free variables = x, v

- Choose Entering variables = x

  Leaving variables = w

- Replace w with x

Basic	Z	x	y	w	v	solution	Intercept
Z	1	0	-1/2	-1/2	0	3	
x	0	1	½	-1/2	0	3	6
y	0	0	3/2	1/2	-1	3	2

Entering variables = y

Leaving variables = y

Basic	Z	x	y	w	v	solution	Intercept
Z	1	0	0	-1/3	-1/3	4	
x	0	1	0	-2/3	1/3	2	
y	0	0	1	1/3	-2/3	2	

Optimal corner is at $x=2, y=2$
The minimum is $Z=4$

Gas production problem

Total profit $z = 150x_1 + 175x_2$

Subject to $7x_1 + 11x_2 \le 77$ (Material constraint)

$11x_1 + 8x_2 \le 80$ (time constraint)

$x_1 \le 9$ (regular storage constraint)

$x_2 \le 6$ (premium storage constraint)

$x_1, x_2 \ge 0$ (positivity constraints)

$$Z - 150x_1 - 175x_2 - 0s_1 - 0s_2 - 0s_3 - 0s_4 = 0$$

$$7x_1 + 11x_2 + s_1 = 77$$

$$10x_1 + 8x_2 + s_2 = 80$$

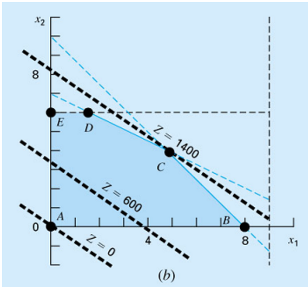
$$x_1 + s_3 = 9$$

$$x_2 + s_4 = 6$$

$Z_{max} = 1400$

$x_1 = 4.9$

$x_2 = 3.9$



The Tableua

Starting point $x_1=0$ and $x_2=0$

Basic	Z	x_1	x_2	s_1	s_2	s_3	s_4	Solution	Intercept
Z	1								
s_1	0								
s_2	0								
s_3	0								
s_4	0								

Basic variables = s_1, s_2, s_3, s_4

Free variables = x_1, x_2

Entering variables = x_1

Leaving variables = s_2

Basic	Z	x_1	x_2	s_1	s_2	s_3	s_4	Solution	Intercept
Z	1	-150	-175	0	0	0	0	0	
s_1	0	7	11	1	0	0	0	77	
x_1	0	1	8/10	0	1/10	0	0	8	
s_3	0	1	0	0	1	1	0	9	
s_4	0	0	1	0	0	0	1	6	

Apply Gauss-Jordan method

Basic	Z	x_1	x_2	s_1	s_2	s_3	s_4	Solution	Intercept
Z	1	0	-55	0	15	0	0	1200	
s_1	0	0	5.4	1	-0.7	0	0	21	21/5.4
x_1	0	1	0.8	0	0.1	0	0	8	10
s_3	0	0	-0.8	0	-0.1	1	0	1	-1.25
s_4	0	0	1	0	0	0	1	6	6

The 2nd corner is at $x_1 = 8, x_2 = 0$

Entering variables = x_2

Leaving variables = s_1

Basic	Z	x_1	x_2	s_1	s_2	s_3	s_4	Solution	Intercept
Z	1	0	0	10.1852	7.8704	0	0	1413.889	
x_2	0	0	1	0.1852	-0.1296	0	0	3.889	
x_1	0	1	0	-0.1481	0.2037	0	0	4.889	
s_3	0	0	0	0.1481	-0.2037	1	0	4.11	
s_4	0	0	0	-0.1852	0.1296	0	1	2.11	

The 3rd corner is at $x_1 = 4.889, x_2 = 3.889$

The maximum is $Z=1413.889$