Mathematics for Economists

Chapters 4-5 – Part 2

Linear Models and Matrix Algebra

Pierre-Simon Laplace (1749–1827, France)

A. L. Cauchy (1789–1857, France)

4.7 Determinant of a Matrix

- The *determinant* is a number associated with any squared matrix.
- If **A** is an *n*x*n* matrix, the determinant is $|\mathbf{A}|$ or det(**A**).
- Since the early days, a determinant was used to "*determine*" if a system of linear equations has a unique solution.
- Cramer (1750) expanded the concept to sets of equations, but a bit later, they were recognized by Vandermonde (1772) as independent functions.
- Determinants are used to characterize invertible matrices. A matrix is invertible (non-singular) if and only if $|\mathbf{A}|\neq 0$.
- That is, if $|\mathbf{A}|\neq 0 \rightarrow \mathbf{A}$ is invertible or non-singular.
- Can be found using factorials, pivots, and cofactors!
- Lots of interpretations.

2

4.7 Determinant of a Matrix

- When n is small, determinants are used for inversion and to solve systems of equations.

Example: Inverse of a 2x2 matrix:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad |A| = \det(A) = ad - bc$$

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

This matrix is called the adjugate of A (or adj(A)).

$A^{-1} = \mathrm{adj}(A)/|A|$

4.7 Determinant of a Matrix (3x3)

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - afh - bdi - ceg$$

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

Sarrus' Rule: Sum from left to right. Then, subtract from right to left

Note: N! terms

- Q: How many flops? For **A** (*3*x3*)*, we count 17 operations.

4.7 Determinants: Laplace formula

- The determinant of a matrix of arbitrary size can be defined by the *Leibniz formula* or the *Laplace formula*.
- The *Laplace formula* (or *expansion*) expresses the determinant $|\boldsymbol{A}|$ as a sum of n determinants of $(n-1) \times (n-1)$ sub-matrices of **A**. There are n^2 such expressions, one for each row and column of **A**
- Define the *i,j minor* M_{ij} (usually written as $|M_{ij}|$) of **A** as the determinant of the $(n-1) \times (n-1)$ matrix that results from deleting the i-th row and the j-th column of **A**.

Pierre-Simon Laplace (1749 – 1827, France).

4.7 Determinants: Laplace formula

- Define the $C_{i,j}$ the *cofactor* of **A** as:

$$C_{i,j} = (-1)^{i+j} \, | M_{i,j} |$$

- The cofactor matrix of **A** -denoted by **C**-, is defined as the nxn matrix whose (i,j) entry is the (i,j) cofactor of **A**. The transpose of **C** is called the adjugate or adjoint of **A** -adj(**A**).

- **Theorem** (Determinant as a Laplace expansion)
  Suppose **A** = $[a_{ij}]$ is an nxn matrix and $i,j= \{1, 2, ...,n\}$. Then the determinant

$$\begin{aligned} |A| &= a_{i1}C_{i1} + a_{i2}C_{i2} + \ldots + a_{in}C_{in} \\ &= a_{ij}C_{ij} + a_{2j}C_{2j} + \ldots + a_{nj}C_{nj} \end{aligned}$$

6

4.7 Determinants: Laplace formula

Example:

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

- $|\mathbf{A}| = 1 * C_{11} + 2 * C_{12} + 3\ C_{13}$ (expanding through 1st row)
  $= 1*(-1)^{1+1}*(-1*6) + 2*(-1)^{1+2}*(0) + 3*(-1)^{1+3}*[-(-1)*2)] = 0$
  $= 2 * C_{12} + (-1) * C_{22} + 4 * C_{23}$ (expanding through 2nd col)
  $= 2*(-1)^{1+2}*(0) + (-1)*(-1)^{2+2}*(0) + 4*(-1)^{2+3}*(0) = 0$

- $|\mathbf{A}|=0$ ⇒ The matrix is singular. (Check!)

• How many flops? For a **A** (3x3), we count 14 operations (better!). For **A** (nxn), we calculate n subdeterminants, each of which requires (n-1) subdeterminants, etc. Then, computations of order n! (plus some n terms), or O(n!).

7

4.7 Determinants: Computations

• By today's standards, a 30×30 matrix is small. Yet it would be impossible to calculate a 30×30 determinant by Laplace formula. It would require over n! (30! ≈ 2.65×10^{32}) multiplications.

• If a computer performs one quatrillion ($1.0\text{x}10^{15}$) multiplications per second (a Petaflops, the 2008 record), it would have to run for over 8.4 billion years to compute a 30×30 determinant by Laplace's method.

• Using a very fast computer like the 2013 China Tianhe-2 (33 petaflops), it would take 254 million years.

• Not a very useful, computationally speaking, method. Avoid factorials! There are more efficient methods.

8

4.7 Determinants: Computations

- Faster way of evaluating the determinant: Bring the matrix to UT (or LT) form by linear transformations. Then, the determinant is equal to the product of the diagonal elements.

- For **A** (nxn), each linear transformation involves adding a multiple of one row to another row, that is, n or fewer additions and n or fewer multiplications. Since there are n rows, this is a procedure of order n^3 -or $O(n^3)$.

Example: For n = 30, we go from 30! = 2.65*10^{32} flops to 30^3 = 27,000 flops.

9

4.7 Determinants: Properties

- Interchange of rows and columns does not affect $|\mathbf{A}|$. (Corollary, $|\mathbf{A}| = |\mathbf{A}'|$.)
- To any row (column) of $\boldsymbol{A}$ we can add any multiple of any other row (column) without changing $|\mathbf{A}|$.

(Corollary: if we transform **A** into **U** or **L**, $|\mathbf{A}|=|\mathbf{U}| = |\mathbf{L}|$, which is equal to the product of the diagonal element of **U** or **L**.)

- $|\mathbf{I}| = 1$, where **I** is the identity matrix.
- $|k\mathbf{A}| = k^n\,|\mathbf{A}|$, where k is a scalar.
- $|\mathbf{A}| = |\mathbf{A}'|$.
- $|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$.
- $|\mathbf{A}^{-1}|=1/|\mathbf{A}|$.
- Recursive flops formula: $\text{flops}_n = n * (\text{flops}_{n-1} + 2) - 1$

10

4.7 Determinants: R

- Simple command, det(A)
- > M = cbind(rbind(1,2), rbind(6,5))

```
     [,1] [,2]
[1,]    1    6
[2,]    2    5
>det(M)
[1] -7
> det(M*2)
[1] -28
> Minv <-solve(M); Minv
           [,1]       [,2]
[1,] -0.7142857  0.8571429
[2,]  0.2857143 -0.1428571
> det(Minv)                    # =1/det(M)
[1] -0.1428571
```

11

4.7 Determinants: Cramer's Rule - Derivation

- Recall the solution to **Ax** = ***d***, where **A** is an *nxn* matrix:

  x* = $\mathbf{A}^{-1}\boldsymbol{d}$

Using the cofactor method to get the inverse we get:

$$x^* = \frac{1}{|A|}(\text{adjoint } A) \quad (d)$$

Note: **A**$|\neq 0 \rightarrow$ **A** is non-singular.

$$\underset{nx1}{\begin{bmatrix} x_1^* \\ x_2^* \\ \dots \\ x_n^* \end{bmatrix}} = \underset{1x1}{\frac{1}{|A|}} \underset{nxn}{\begin{bmatrix} |C_{11}| & |C_{21}| & \cdots & |C_{n1}| \\ |C_{12}| & |C_{22}| & \cdots & |C_{n2}| \\ \dots & \dots & & \dots \\ |C_{1n}| & |C_{2n}| & \cdots & |C_{nn}| \end{bmatrix}} \underset{nx1}{\begin{bmatrix} d_1 \\ d_2 \\ \dots \\ d_n \end{bmatrix}} = \frac{1}{|A|} \begin{bmatrix} \sum_{i=1}^{n} d_i |C_{i1}| \\ \sum_{i=1}^{n} d_i |C_{i2}| \\ \dots \\ \sum_{i=1}^{n} d_i |C_{in}| \end{bmatrix}$$

12

4.7 Determinants: Cramer's Rule - Derivation

- **Example**: Let **A** be 3x3. Then,

1) $$\begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} d_1|C_{11}|+d_2|C_{21}|+d_3|C_{31}| \\ d_1|C_{12}|+d_2|C_{22}|+d_3|C_{32}| \\ d_1|C_{13}|+d_2|C_{23}|+d_3|C_{33}| \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} \sum_{i=1}^{3} d_i|C_{i1}| \\ \sum_{i=1}^{3} d_i|C_{i2}| \\ \sum_{i=1}^{3} d_i|C_{i3}| \end{bmatrix}$$

2) $$\sum_{i=1}^{3} d_i|C_{i1}| = d_1|C_{11}|+d_2|C_{21}|+d_i|C_{31}| \quad \text{where } |C_{ij}| \equiv (-1)^{i+j}|M_{ij}|$$

3) $$\sum_{i=1}^{3} d_i|C_{i1}| = d_1\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + d_2\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + d_3\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = |A_1|$$

4) $$A_1 = \begin{bmatrix} d_1 & a_{12} & a_{13} \\ d_2 & a_{22} & a_{23} \\ d_3 & a_{32} & a_{33} \end{bmatrix}. \quad \text{Find } |A_1| \text{ such that } x_1^* = |A_1|/|A|$$

13

4.7 Determinants: Cramer's Rule - Derivation

$$x_1^* = \frac{\sum_{i=1}^{3} d_i|C_{i1}|}{\sum_{i=1}^{3} a_{i1}|C_{i1}|} = \frac{\begin{vmatrix} d_1 & a_{12} & a_{13} \\ d_2 & a_{22} & a_{23} \\ d_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} = \frac{|A_1|}{|A|}$$

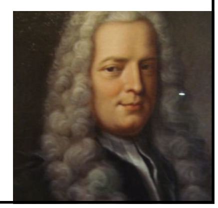
$$x_2^* = \frac{\sum_{i=1}^{3} d_i|C_{i2}|}{\sum_{i=1}^{3} a_{i2}|C_{i2}|} = \frac{\begin{vmatrix} a_{11} & d_1 & a_{13} \\ a_{21} & d_2 & a_{23} \\ a_{31} & d_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} = \frac{|A_2|}{|A|}$$

14

4.7 Determinants: Cramer's Rule - Derivation

$$x_3^* = \frac{\sum_{i=1}^{3} d_i |C_{i3}|}{\sum_{i=1}^{3} a_{i3} |C_{i3}|} = \frac{\begin{vmatrix} a_{11} & a_{12} & d_1 \\ a_{21} & a_{22} & d_2 \\ a_{31} & a_{32} & d_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} = \frac{|A_3|}{|A|}$$

Gabriel Cramer (1704-1752, Switzerland).



4.7 Determinants: Cramer's Rule - Derivation

- Following the pattern, we have the general Cramer's rule:

$$\begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \\ \ldots \\ x_n^* \end{bmatrix} = \frac{1}{|A|} \begin{bmatrix} \sum_{i=1}^{n} d_i |C_{i1}| \\ \sum_{i=1}^{n} d_i |C_{i2}| \\ \sum_{i=1}^{n} d_i |C_{i3}| \\ \ldots \\ \sum_{i=1}^{n} d_i |C_{in}| \end{bmatrix} = \begin{bmatrix} |A_1|/|A| \\ |A_2|/|A| \\ |A_3|/|A| \\ \ldots \\ |A_n|/|A| \end{bmatrix}$$

16

4.7 Cramer's Rule Application: Macro Model

Matrix form

$$\begin{bmatrix} 1 & -1 & -1 \\ -b & 1 & 0 \\ -g & 0 & 1 \end{bmatrix}\begin{bmatrix} Y \\ C \\ G \end{bmatrix} = \begin{bmatrix} I_0 \\ a - bT_0 \\ 0 \end{bmatrix}$$

The determinant

$$|A| = \begin{vmatrix} 1 & -1 & -1 \\ -b & 1 & 0 \\ -g & 0 & 1 \end{vmatrix} = 1 - (b+g)$$

17

4.7 Cramer's Rule Application: Macro Model

- Applying Cramer's rule for the 3x3 case:

$$|A_Y| = \begin{vmatrix} I_0 & -1 & -1 \\ a-bT_0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = I_0 + a - bT_0 \qquad Y^* = \frac{|A_Y|}{|A|} = \frac{I_0 + a - bT_0}{1-(b+g)}$$

$$|A_C| = \begin{vmatrix} 1 & I_0 & -1 \\ -b & a-bT_0 & 0 \\ -g & 0 & 1 \end{vmatrix} = bI_0 + (1-g)(a-bT_0) \qquad C^* = \frac{|A_C|}{|A|} = \frac{bI_0 + (1-g)(a-bT_0)}{1-(b+g)}$$

$$|A_G| = \begin{vmatrix} 1 & -1 & I_0 \\ -b & 1 & a-bT_0 \\ -g & 0 & 0 \end{vmatrix} = g(a - bT_0 + I_0) \qquad G^* = \frac{|A_G|}{|A|} = \frac{g(a - bT_0 + I_0)}{1-(b+g)}$$

18

4.8 - Linear Models & Matrix Algebra: Summary

- Matrix algebra can be used:

a. to express the system of equations in a compact notation;

b. to find out whether solution to a system of equations exist; and

c. to obtain the solution if it exists.

d. If n is small, we can find $\mathbf{A}^{-1}$, but in general, we will avoid this step. We will resort to more efficient methods to solve for x*, likely using a Cholesky decomposition with Gaussian elimination.

$$Ax = d$$

$$x^* = A^{-1}d$$

$$A^{-1} = \frac{adjA}{\det A}$$

$$x^* = \frac{adjA}{|A|}d$$

19

4.8 - Notation and Definitions: Summary

- **A** (Upper case letters) = matrix
- b (Lower case letters) = vector
- nxm = n rows, m columns
- rank(**A**) = number of linearly independent vectors of **A**
- trace(**A**) = tr(**A**) = sum of diagonal elements of **A**
- Null matrix = all elements equal to zero.
- Diagonal matrix = all off-diagonal elements are zero.
- **I** = identity matrix (diagonal elements: 1, off-diagonal: 0)
- $|\mathbf{A}|$ = det(**A**) = determinant of **A**
- $\mathbf{A}^{-1}$ = inverse of **A**
- **A**'= $\mathbf{A}^T$ = Transpose of **A**
- $|M_{ij}|$ = Minor of **A**
- $\mathbf{A} = \mathbf{A}^T$ ⇒ Symmetric matrix
- $\mathbf{A}^T\mathbf{A} = \mathbf{A}\mathbf{A}^T$ ⇒ Normal matrix
- $\mathbf{A}^T = \mathbf{A}^{-1}$ ⇒ Orthogonal matrix
- $\mathbf{A} = \mathbf{A}^2$ ⇒ Idempotent matrix

20

4.9 Eigenvalues and Diagonal Systems

- A set of linear simultaneous equations:

  Ab = **d** **A** is a non-singular and **b** and **d** are conformable vectors.

- Under certain circumstances, we can diagonalize this system:

  $\boldsymbol{\Lambda\nu} = \upsilon$ where $\boldsymbol{\Lambda}$ is a diagonal matrix.

- Eigenvalues (Characteristic Roots)

  Ax = λ**x** This is the eigenvalue problem
  λ is the *eigenvalue* (*characteristic root*)
  x is the *eigenvector* (*characteristic vector*)

- Cauchy discovered them studying how to find new coordinate axes for the graph of the quadratic equation $ax^2+2bxy+cy^2=d$ so that the equation with the new axes would be of the form $Ax^2+Cy^2 = D$.

21

4.9 Eigenvalues and Diagonal Systems

- **Ax** = λ **x** (Basic equation of eigenvalue problem)

- For the square matrix **A**, there is a vector **x** such that the product of **Ax** results in a scalar, λ, that, when multiplied by **x**, gives the same product.

- The multiplication of vector **x** by a scalar is the same as stretching or shrinking the coordinates by a constant value. (The matrix **A** just scales the vector x!)

- **Ax** = λ **Ix** ⇒ [**A** - λ **I**] **x** = **0**
- **K** = [**A** - λ **I**] Characteristic matrix of matrix **A**
- **K x** = **0** Homogeneous equations.

22

4.9 Eigenvalues and Diagonal Systems

• Homogeneous equations: $\mathbf{K}\,\mathbf{x} = \mathbf{0}$
  - Trivial solution $\mathbf{x} = \mathbf{0}$ (If $|\mathbf{K}| \neq 0$, from *Cramer's rule*)
  - Nontrivial solution ($\mathbf{x} \neq \mathbf{0}$) can occur if $|\mathbf{K}| = 0$.

• That is, do all matrices have eigenvalues?
No. They must be square and $|\mathbf{K}| = |\mathbf{A} - \lambda\mathbf{I}| = 0$.

• Eigenvectors are not unique. If $\mathbf{x}$ is an eigenvector, then $\beta\mathbf{x}$ is also an eigenvector: $\mathbf{A}(\beta\mathbf{x}) = \lambda\,(\beta\mathbf{x})$

• To calculate eigenvectors and eigenvalues, expand the equation
$$|\mathbf{A} - \lambda\mathbf{I}| = 0$$

• The resulting equation is called *characteristic equation*.

23

4.9 Eigenvalues and Diagonal Systems

• Characteristic equation: $|\mathbf{A} - \lambda\mathbf{I}| = 0$

Example: For a 2x2 matrix:

$$[\mathrm{A} - \lambda \mathrm{I}] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix}$$

$$|\mathrm{A} - \lambda \mathrm{I}| = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0$$

$$a_{11}a_{22} - a_{12}a_{21} - \lambda(a_{11} + a_{22}) + \lambda^2 = 0$$

• For a 2-dimensional problem, we have a simple quadratic equation with two solutions for λ.

24

4.9 Eigenvalues: 2x2 Case

For n=2, we have a simple quadratic equation with two solutions for λ. In fact, there is generally one eigenvalue for each dimension, but some may be zero, and some complex.

$$0 = a_{11}a_{22} - a_{12}a_{21} - (a_{11} + a_{22})\lambda + \lambda^2$$

$$\lambda = \frac{(a_{11} + a_{22}) \pm \sqrt{(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}}{2}$$

Note 1: The solution for λ can be written as:

$\lambda = ½\text{trace}(A) \pm ½\,[\text{trace}(A)^2 - 4|A|]^{1/2}$

Three cases:

1) Real different roots: $\text{trace}(A)^2 > 4|A|$
2) One real root: $\text{trace}(A)^2 = 4|A|$
3) Complex roots: $\text{trace}(A)^2 < 4|A|$

4.9 Eigenvalues: 2x2 Case

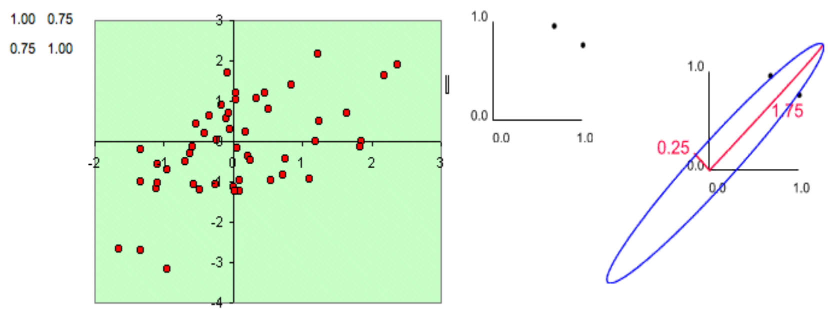
• Note 2: If **A** is symmetric, the eigenvalues are real. That is, we need to have trace(**A**)$^2 > 4|$**A**$|$. For n=2, we check this condition:

$$(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{12}) > 0$$
$$a_{11}^2 + a_{22}^2 + 2a_{11}a_{22} - 4a_{11}a_{22} + 4a_{12}^2 > 0$$
$$a_{11}^2 + a_{22}^2 - 2a_{11}a_{22} + 4a_{12}^2 > 0$$
$$(a_{11} - a_{22})^2 + 4a_{12}^2 > 0$$

4.9 Eigenvalues: Geometric Interpretation

- Geometric Interpretation for the 2x2 Case:

The *x,y* values of A can be seen as representing points on an ellipse centered at (0,0). The eigenvectors are in the directions of the major and minor axes of the ellipse, and the eigenvalues are the lengths of these axes to the ellipse from (0,0).



4.9 Eigenvalues: General Case

- General *n*x*n* case:

The characteristic determinant $D(\lambda) = \det(\mathbf{A} - \lambda\,\mathbf{I})$
is clearly a polynomial in λ:

$$D(\lambda) = \alpha_n \lambda^n + \alpha_{n-1} \lambda^{n-1} + \alpha_{n-2} \lambda^{n-2} + \ldots + \alpha_1 \lambda + \alpha_0$$

- Characteristic equation:

$$D(\lambda) = \alpha_n \lambda^n + \alpha_{n-1} \lambda^{n-1} + \alpha_{n-2} \lambda^{n-2} + \ldots + \alpha_1 \lambda + \alpha_0 = 0$$

There are *n* solutions to this polynomial. The set of eigenvalues is called the *spectrum* of **A**. The largest of the absolute values of the eigenvalues of **A** is called the *spectral radius* of **A.**

- Eigenvalues are computed using the *QR algorithm* (1950s) or the *divide-and-conquer eigenvalue algorithm* (1990s). They are computationally intensive. They take 4 $n^3/3$ flops.

4.9 Eigenvalues: Properties

- Some properties:
  - The product of the eigenvalues = $|\mathbf{A}|$
  - The sum of the eigenvalues = trace($\mathbf{A}$)
  - The eigenvalues of $\mathbf{A}^k$ are $\lambda_1^k, \lambda_2^k, ..., \lambda_n^k$.
  - If $\mathbf{A}$ is an idempotent matrix, its λ_i's are all 0 or 1.
  - If $\mathbf{A}$ is an orthogonal matrix ($\mathbf{A}^T = \mathbf{A}^{-1}$), its λ_i's (if real) are ±1.
    Proof: $\lambda^2\, \mathbf{x'x} = \mathbf{x'A'Ax} = \mathbf{x'x} \Rightarrow |\lambda| = 1$ (if real, $\lambda = \pm 1$).
  - If $\mathbf{A}$ is a symmetric (Hermitian) matrix:
    - its λ_i's are all real.
    - its eigenvectors are orthogonal.
  - All eigenvectors derived from unequal eigenvalues are linearly independent. (*n* eigenvectors can form an orthonormal basis!).
  - If $\mathbf{A}$ is a pd matrix, its λ_i's are positive and, then, $|\mathbf{A}|>0$.
    Proof: $0 < \mathbf{x'Ax} = \mathbf{x'}\lambda\mathbf{x} = \lambda\mathbf{x'x} = \lambda\|\mathbf{x}\|^2$.
    Since $\|\mathbf{x}\|^2>0 \quad \Rightarrow \lambda>0$ (&$|\mathbf{A}|>0$).

4.9 Eigenvalues: Example -Correlation Matrix

- **Example**: A correlation matrix

$$\mathbf{A} = \begin{bmatrix} 1 & .75 \\ .75 & 1 \end{bmatrix}$$

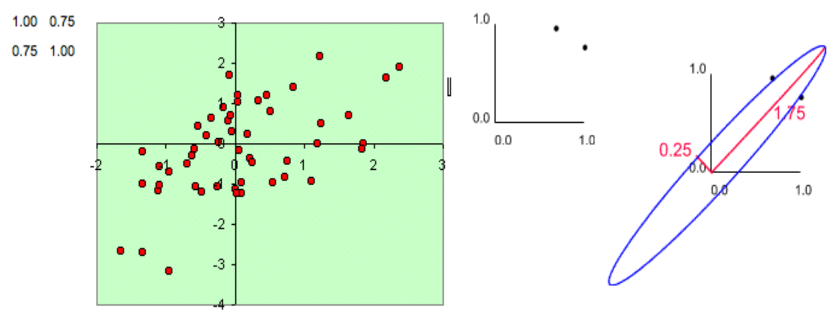
$$\begin{aligned}\lambda &= ½\text{trace}(\mathbf{A}) \pm ½\, [\text{trace}(\mathbf{A})^2 - 4\,|\mathbf{A}|]^{1/2} \\ &= ½\, 2 \pm ½[2^2 - 4*0.4735]^{1/2} = 1 \pm ½[2.25]^{1/2} \\ &= 1 \pm ½[1.5] = 0.25;\ 1.75\end{aligned}$$

x = [-0.7071,   0.7071]; [0.7071,   0.7071]

Note: $\mathbf{x}$ is not unique. Usually, we set $\|\mathbf{x}\|$=1 (dot product).

4.9 Eigenvalues: Example -Correlation Matrix

- Graphical interpretation: Correlation as an ellipse, whose major axis is one eigenvalue and the minor axis length is the other:



No correlation yields a circle, and perfect correlation yields a line.

4.9 Eigenvalues: R Commands

- Command “eigen,” recover values with $

```
> A <- matrix(c(1, .75, .75, 1), nrow = 2)
> A
     [,1] [,2]
[1,] 1.00 0.75
[2,] 0.75 1.00
> eigen(A)
eigen() decomposition
$values
[1] 1.75 0.25                    ⇒ positive real eigenvalues, A is pd!

$vectors
          [,1]       [,2]
[1,] 0.7071068 -0.7071068
[2,] 0.7071068  0.7071068.       ⇒ symmetric matrix, eigenvalues are orthogonal
> lamb <-eigen(A)
> lambda <- lamb$values
> lambda
[1] 1.75 0.25
```

4.9 Eigenvalues: R Commands

- Command "eigen," recover values with $

```
> x_lamb <- lamb$vectors
> x_lamb[1]
[1] 0.7071068
> x_lamb[1,]
[1]  0.7071068 -0.7071068
> x_lamb[2,]
[1] 0.7071068 0.7071068
> t(x_lamb[1,])%*%x_lamb[2,]
     [,1]
[1,]    0                                   ⇒ yes, eigenvectors are orthogonal.
>
```

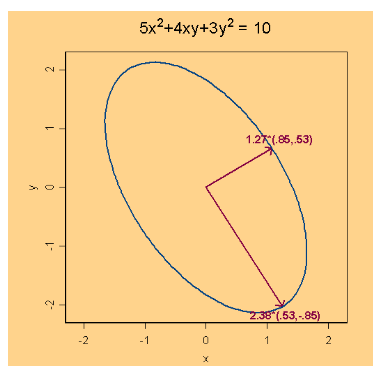
4.9 Eigenvalues: Example – Quadratic Form

- 2^{nd} order multivariable equations: $ax^2 + 2\,kxy + by^2 = c$
- Represented in a quadratic form with symmetric matrix A:

$$\mathbf{x}^T \mathbf{A}\, \mathbf{x} = \mathbf{c}, \quad \text{where}$$

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} a & k \\ k & b \end{pmatrix}$$

$$5x^2 + 4xy + 3y^2 = 10 \qquad \mathbf{A} = \begin{pmatrix} 5 & 2 \\ 2 & 3 \end{pmatrix}$$



Eigenvector decomposition:

$\lambda_1 = 1.764$, $x_1 = [.5257,-.8507]$

$\lambda_2 = 6.236$, $x_2 = [-.8507,-.5257]$

- Positive Definite **A** ⇒ positive real eigenvalues!
- Symmetric **A** ⇒ <u>*orthogonal e-vectors*</u>!
- Geometrical interpretation: <u>Principal Axes of Ellipse.</u>

4.9 Diagonal (Eigen) decomposition

- Let **A** be a square nxn matrix with n linearly independent eigenvectors, x_i (i = 1, 2, …, n). Then, **A** can be factorized as

$$\mathbf{A} = \mathbf{X}\,\boldsymbol{\Lambda}\,\mathbf{X}^{-1}$$

where **X** is the square (nxn) matrix whose i^{th} column is the eigenvector x_i of **A** and **Λ** is the diagonal matrix whose diagonal elements are the corresponding eigenvalues, *i.e.*, $\Lambda_{ii} = \lambda_i$.

- The eigenvectors are usually normalized, but they need not be. A non-normalized set of eigenvectors can also be used as the columns of **X**.

Proof: $\mathbf{Ax} = \lambda\mathbf{x}$ $\Rightarrow \mathbf{AX}=\mathbf{X}\boldsymbol{\Lambda}$ $\Rightarrow \mathbf{A} = \mathbf{X}\boldsymbol{\Lambda}\mathbf{X}^{-1}$ ($\mathbf{X}^{-1}$ exists)■

- Conversely: $\mathbf{X}^{-1}\,\mathbf{A}\,\mathbf{X} = \boldsymbol{\Lambda}$
- If $\mathbf{X}^T\,\mathbf{X} = \mathbf{I}$, **A** is *orthogonally* diagonalizable.

4.9 Diagonal decomposition: Example

- Let $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}; \lambda_1 = 1, \lambda_2 = 3.$

The eigenvectors: $\mathbf{x}_1$=[1,-1], $\mathbf{x}_2$=[1,1].

- Let **X** be the matrix of eigenvectors: $\mathbf{X} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$
- Inverting, we have $\mathbf{X}^{-1} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}$
- Then, $\boldsymbol{A} = \mathbf{X}\,\boldsymbol{\Lambda}\,\mathbf{X}^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}$

4.9 Diagonal decomposition: Example

- Diagonalizing a system of equations:

  $\mathbf{A\,x} = \mathbf{y}$
- Pre-multiply both sides by $\mathbf{X}^{-1}$:

  $\mathbf{X}^{-1}\,\mathbf{A\,x} = \mathbf{X}^{-1}\,\mathbf{y} = \boldsymbol{\nu}$

  $\mathbf{X}^{-1}\,\mathbf{A}\,(\mathbf{X}\,\mathbf{X}^{-1})\mathbf{x} = \boldsymbol{\nu}$ (Let $\boldsymbol{\upsilon} = \mathbf{X}^{-1}\,\mathbf{x}$)

  $\Rightarrow \boldsymbol{\Lambda}\,\boldsymbol{\upsilon} = \boldsymbol{\nu}$
- Using the (2x2) previous example:

$$\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$\upsilon_1 = \frac{1}{2}x_1 - \frac{1}{2}x_2$ $\frac{1}{2}y_1 - \frac{1}{2}y_2 = \nu_1$

$\upsilon_2 = \frac{1}{2}x_1 + \frac{1}{2}x_2$ $\frac{1}{2}y_1 + \frac{1}{2}y_2 = \nu_2$

$\upsilon_1 = \nu_1$

$3\upsilon_2 = \nu_2$

4.9 Diagonal decomposition: Application

- Let $\mathbf{M}$ be the square nxn matrix defined by: $\mathbf{M} = \mathbf{I}_n - \mathbf{Z}(\mathbf{Z}'\mathbf{Z}^{-1})\mathbf{Z}'$, where $\mathbf{Z}$ is an nxk matrix, with rank($\mathbf{Z}$)=k .

Let's calculate the trace($\mathbf{M}$):

trace($\mathbf{M}$) = $\text{tr}(\mathbf{I}_n - \mathbf{Z}(\mathbf{Z}'\mathbf{Z}^{-1})\mathbf{Z}') = \text{tr}(\mathbf{I}_n) - \text{tr}(\mathbf{Z}(\mathbf{Z}'\mathbf{Z}^{-1})\mathbf{Z}') =$

$= n - \text{tr}((\mathbf{Z}'\mathbf{Z}^{-1})\mathbf{Z}'\mathbf{Z}) = n - \text{tr}(\mathbf{I}_k) = n - k.$

It is easy to check that $\mathbf{M}$ is idempotent (λ_i's are all 0 or 1) and symmetric (λ_i's are all real and $\mathbf{x}$ are orthogonal).

Write an orthogonal diagonalization: $\mathbf{M} = \mathbf{X}\,\boldsymbol{\Lambda}\,\mathbf{X}^{-1}$ ($\mathbf{X}'\mathbf{X}^{-1} = \mathbf{I}$).

Again, let's calculate the trace($\mathbf{M} = \mathbf{X}\,\boldsymbol{\Lambda}\,\mathbf{X}^{-1}$):

trace($\mathbf{M}$) = $\text{tr}(\mathbf{X}\boldsymbol{\Lambda}\mathbf{X}^{-1}) = \text{tr}(\boldsymbol{\Lambda}\mathbf{X}^{-1}\mathbf{X}) = \text{tr}(\boldsymbol{\Lambda}) = \Sigma_i\,\lambda_i$

That is, $\mathbf{M}$ has $n - k$ non-zero eigenvalues.

4.9 Why do Eigenvalues/vectors matter?

- Eigenvectors are invariants of **A**
  - Don't change direction when operated **A**
- Use to determine the definiteness of a matrix.
- A singular matrix has at least one zero eigenvalue.
- Solutions of multivariable differential equations (the bread-and-butter in linear systems) correspond to solutions of linear algebraic eigenvalue equations.
- Eigenvalues are used to study the stability of autoregressive time series models.
- The orthogonal basis of eigenvectors forms the core of principal components analysis (PCA).

4.9 Sign of a quadratic form: Eigenvalue tests

- Suppose we are interesting in an optimization problem for z=$f(x,y)$. We set the first order conditions (f.o.c.), solve for x^* and y^*, and, then, check the second order conditions (s.o.c.).
- Let's re-write the s.o.c. of $z = f(x,y)$:

$$d^2z = q = f_{xx}dx^2 + 2f_{xy}dxdy + f_{yy}dy^2$$

$$q = \begin{bmatrix} dx & dy \end{bmatrix} \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix} = u'Hu$$

- The s.o.c. of z=$f(x,y)$ is a quadratic form, with a symmetric matrix, **H**.
- To determine what type of extreme points we have, we need to check the sign of the quadratic form.

40

4.9 Sign of a quadratic form: Eigenvalue tests

- .Quadratic form:

  q = u'**H** u (note: the Hessian, **H**, is a symmetric matrix)
- Let u=**T**y, where **T** is the matrix of eigenvectors of **H**, such that

  T' T = I (**H** is symmetric matrix ⇒ **T**'=$\mathbf{T}^{-1}$. Check!)
- . Then,

  q = y' **T' H T** y = y'**Λ** y (**T' H T = Λ**)

  $q = \lambda_1 y_1^2 + \lambda_2 y_2^2 + ... + \lambda_i y_i^2 + ... + \lambda_n y_n^2$

  ⇒ sign(q) depends on the λ_i's only.
- We say:

  q is positive definite iff $\lambda_i > 0$ for all i.

  q is positive semi-definite iff $\lambda_i \geq 0$ for all i.

  q is negative semi-definite iff $\lambda_i \leq 0$ for all i.

  q is negative definite iff $\lambda_i < 0$ for all i.

  q is indefinite if some $\lambda_i > 0$ and some $\lambda_i < 0$.

41

4.9 Sign of a quadratic form: Eigenvalue tests

- **Example**: Find extreme values for z=*f*(*x*,*y*), and determine if they are a max or min.

$$z = x^2 + xy + 2y^2 + 3$$

F.o.c.

$$\begin{cases} f_x = 2x + y = 0 \\ f_y = x + 4y = 0 \end{cases}$$

$$y^* = 0,\ x^* = 0,\ z^* = 3$$

Calculate matrix of second derivatives

$$|H| = \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} = \begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix}, \qquad \lambda_1 = 1.58582; \quad \lambda_2 = 4.4142$$

=> λ_1 and λ_2 are positive, q is positive definite

=> z^* is minimum

42

4.10 Linear Algebra: Application

Use linear algebra to find the identity of superman.