R403: Probabilistic and Statistical Computations with R

Topic 2: Matrix Algebra. Eigenvalues and Eigenvectors. R's Matrix Algebra Capabilities. Complex Numbers

Kaloyan Ganev

2022/2023

Lecture Contents

Matrix Algebra Basics

Matrix Algebra Basics

- An $m \times n$ matrix $\mathbf{A}$ is a rectangular array of numbers:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

- How to create one in R (say $m = 3, n = 4$):

```r
A <- matrix(1:12, nrow = 3)
```

- Another one (we'll need it in a while):

```r
B <- matrix(13:24, nrow = 3)
```

Matrix Algebra Basics (2)

- Matrix addition and subtraction:

$$\mathbf{C} = \mathbf{A} + \mathbf{B} \Leftrightarrow c_{ij} = a_{ij} + b_{ij}$$
$$\mathbf{D} = \mathbf{A} - \mathbf{B} \Leftrightarrow d_{ij} = a_{ij} - b_{ij}$$

Note: $\mathbf{A}$ and $\mathbf{B}$ must be conformable!

- In R:

```
C <- A + B
D <- A - B
```

- Matrix transposition:

$$\mathbf{E} = \mathbf{A}' \Leftrightarrow e_{ij} = a_{ji}$$

```
E <- t(A)
```

Matrix Algebra Basics (3)

- Scalar multiplication:
$$\mathbf{F} = \alpha\mathbf{A} \Leftrightarrow f_{ij} = \alpha a_{ij}$$
- Matrix multiplication: the generic element of the product $\mathbf{G} = \mathbf{AE}$, g_{ij}, equals
$$g_{ij} = \sum_{r=1}^{n} a_{ir} e_{rj}$$

```
G <- A %*% E
```

- Conformity rule: the number of columns in $\mathbf{A}$ must equal the number of rows in $\mathbf{E}$
- Check out also

```
crossprod(A, B)
```

Matrix Algebra Basics (4)

- Associative law:
$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$$
- Left and right distributive laws:
$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$$
$$(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$$
- Note that in the general case $\mathbf{AB} \neq \mathbf{BA}$
- Also, if $\mathbf{A} \neq \mathbf{0}$, $\mathbf{AB} = \mathbf{AC}$ does not imply $\mathbf{B} = \mathbf{C}$
- Square matrix: $m = n$ (the number of rows equals the number of columns)
- For a positive integer n,
$$\mathbf{A}^n = \underbrace{\mathbf{AAA}\ldots\mathbf{A}}_{n \text{ times}}$$

Matrix Algebra Basics (5)

- Diagonal matrix: all off-main-diagonal elements are 0:

$$\mathbf{D} = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}$$

- A convenient feature:

$$\mathbf{D}^n = \begin{pmatrix} d_1^n & 0 & \dots & 0 \\ 0 & d_2^n & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n^n \end{pmatrix}$$

- Note that we referred to diagonal *square* matrices, but sometimes they can have differing numbers of rows and columns (rectangular matrices)

Matrix Algebra Basics (6)

- Identity matrix: special case of a diagonal matrix

$$\mathbf{I}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

- Create a (square) diagonal matrix in R:

```
D <- diag(c(1,2,3))
```

- A 7×7 identity matrix:

```
I <- diag(7)
```

- Note that $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$ for every conformable square matrix $\mathbf{A}$

Matrix Algebra Basics (7)

- Rules for transpose matrices:

$$
\begin{aligned}
(\mathbf{A}')' &= \mathbf{A} \\
(\mathbf{A}+\mathbf{B})' &= \mathbf{A}' + \mathbf{B}' \\
(\alpha\mathbf{A})' &= \alpha\mathbf{A}' \\
(\mathbf{AB})' &= \mathbf{B}'\mathbf{A}'
\end{aligned}
$$

- If $\mathbf{A}$ is a square matrix and $\mathbf{A}' = \mathbf{A}$, then $\mathbf{A}$ is called *symmetric*
- Trace of a square matrix: the sum of all main diagonal elements; to calculate

```
library(psych)
tr(I)
```

- Orthogonal square matrix:

$$\mathbf{AA}' = \mathbf{A}'\mathbf{A} = \mathbf{I} \quad (\Leftrightarrow \mathbf{A}' = \mathbf{A}^{-1})$$

Matrix Algebra Basics (8)

- Determinant: can be computed only for square matrices!
- For 2×2 and 3×3 matrices: recall the simple rules
- For $n \times n$ matrices, the rule is a bit more complex
- Some rules of computation:

$$
\begin{aligned}
|\mathbf{A}'| &= |\mathbf{A}| \\
|\mathbf{AB}| &= |\mathbf{A}| \cdot |\mathbf{B}| \\
|\mathbf{A} + \mathbf{B}| &\neq |\mathbf{A}| + |\mathbf{B}|
\end{aligned}
$$

  (The third relationship can be an equality only in special cases)
- To calculate the determinant of, say, the matrix $\mathbf{G}$ in R:

```
det(G)
```

Matrix Algebra Basics (9)

- If $\mathbf{A}$ is an $n \times n$ square matrix and $\det(\mathbf{A}) \neq 0$, then there is a unique matrix $\mathbf{B}$ called the *inverse* of $\mathbf{A}$ such that

$$\mathbf{B} = \mathbf{A}^{-1} \Leftrightarrow \mathbf{AB} = \mathbf{I}_n \Leftrightarrow \mathbf{BA} = \mathbf{I}_n$$

- The reverse is also true, i.e. if $\mathbf{A}^{-1}$ exists, then $\det(\mathbf{A}) \neq 0$
- If the inverse exists, then it is given by

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \operatorname{adj}(\mathbf{A}),$$

where $\operatorname{adj}(\mathbf{A})$ is the *adjoint matrix* of $\mathbf{A}$

Matrix Algebra Basics (10)

- The adjoint matrix is itself defined as

$$\operatorname{adj}(\mathbf{A}) = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{21} & \cdots & \mathbf{A}_{n1} \\ \mathbf{A}_{12} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{1n} & \mathbf{A}_{2n} & \cdots & \mathbf{A}_{nn} \end{pmatrix},$$

  and each $\mathbf{A}_{ij}$ is the cofactor of the element a_{ij}

- Some rules concerning inverses:

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}; \quad (\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$
$$(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'; \quad (c\mathbf{A})^{-1} = \frac{1}{c}\mathbf{A}^{-1}$$

- To calculate an inverse in R:

```
A <- matrix(rnorm(9), nrow = 3)
solve(A)
```

Cramer's Rule

- Consider a system of n equations in n unknowns, $x_1, x_2, \ldots, x_n$,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n \end{aligned}$$

- The system has a unique solution if and only if (iff) $\det(\mathbf{A}) \neq 0$
- In such a case, the solution is $x_j = |\mathbf{A}_j| / |\mathbf{A}|$ where $\mathbf{A}_j$ is obtained from $\mathbf{A}$ by replacing the jth column with the vector of b's

$$|\mathbf{A}_j| = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$

Cramer's Rule (2)

- If the RHS of the system has only zeros, the system is called *homogeneous*
- Homogeneous systems always have the trivial solution $x_1 = x_2 = \ldots = x_n = 0$
- A homogeneous system has non-trivial solutions iff $\det(\mathbf{A}) = 0$
- Solve non-homogeneous and homogeneous systems of n equations in n unknowns in R:

```
A <- matrix(c(1.5, 7.1, 2.2, 3.3, 9.5, 6.8, 1.9, 5.4, 8.2), nrow
    = 3)
det(A)
b <- c(1.8, 2.1, 3.0)
solve(A, b)

b1 <- c(0,0,0)
solve(A, b1)
```

Vectors

- Here we stick to the mathematical (algebraic) notion, not R's one
- A vector is an ordered, one-dimensional collection of numbers
- Consider two vectors of size n, $\mathbf{a} = (a_1, a_2, \dots, a_n)$ and $\mathbf{b} = (b_1, b_2, \dots, b_n)$
- Inner (dot) product:

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{n} a_i b_i$$

- If you consider the vectors as $n \times 1$ matrices, then the dot product is $\mathbf{a}'\mathbf{b}$
- Inner product of vectors in R:

```
a <- c(1,2,3)
b <- c(2,3,9)
dotp <- t(a) %*% b
```

Vectors (2)

- Some rules:

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$$
$$\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$$
$$(\alpha \mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\alpha \mathbf{b}) = \alpha (\mathbf{a} \cdot \mathbf{b})$$

- Euclidean norm (length of vector):

$$\|\mathbf{a}\| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{\sum_{i=1}^{n} a_i^2}$$

- Cauchy-Schwarz inequality:

$$|\mathbf{a} \cdot \mathbf{b}| \leq \|\mathbf{a}\| \cdot \|\mathbf{b}\|$$

- Triangle inequality:

$$\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$$

Vectors (3)

- Angle θ between two non-zero vectors $\mathbf{a}$ and $\mathbf{b}$:

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{a}\|\cdot\|\mathbf{b}\|}$$

- Orthogonality:

$$\mathbf{a}\perp\mathbf{b} \Leftrightarrow \mathbf{a}\cdot\mathbf{b} = 0$$

- Straight line through $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^n$:

$$\mathbf{x} = t\mathbf{a} + (1-t)\mathbf{b}, \quad t \in \mathbb{R}^1$$

- Hyperplane passing through $\mathbf{a}$ in $\mathbb{R}^n$ and orthogonal to the non-zero vector $\mathbf{p} = (p_1, \ldots, p_n)$:

$$\mathbf{p}\cdot(\mathbf{x}-\mathbf{a}) = 0$$

Linear (In)dependence

Linear (In)dependence

- Consider the column vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ in $\mathbb{R}^m$
- Those vectors are *linearly dependent* if there exist numbers $c_1, c_2, \ldots, c_n$, not all zero, such that

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \ldots + c_n\mathbf{a}_n = \mathbf{0}$$

- If the latter only holds for $c_1 = c_2 = \ldots = c_n = 0$, the vectors are *linearly independent*
- Now consider a linear system of m equations in n unknowns:

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{array} \Leftrightarrow x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n = \mathbf{b}$$

Linear (In)dependence (2)

- Suppose this system of equations has two solutions, $\mathbf{u}' = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v}' = (v_1, v_2, \ldots, v_n)$
- Then $u_1\mathbf{a}_1 + \ldots + u_n\mathbf{a}_n = \mathbf{b}$, and $v_1\mathbf{a}_1 + \ldots + v_n\mathbf{a}_n = \mathbf{b}$
- Subtract the second equation from the first and define $c_1 = u_1 - v_1, \ldots, c_n = u_n - v_n$; then:

$$c_1\mathbf{a}_1 + \ldots + c_n\mathbf{a}_n = \mathbf{0}$$

- This implies that when there is more than one solution, the columns $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ are linearly dependent
- If the columns are linearly independent, then there would be *at most* one solution
- Whether the system has zero or one solution, depends on the values in $\mathbf{b}$

Linear (In)dependence (3)

Theorem 1

The n column vectors of the square matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

are linearly independent iff $|\mathbf{A}| \neq 0.$

Rank of a Matrix

Rank of a Matrix

Definition 1

The rank of a matrix $\mathbf{A}$ is the maximum number of linearly independent column vectors in it.

- Denoted by $r(\mathbf{A})$ or $rank(\mathbf{A})$
- If $\mathbf{A} = \mathbf{0}$, then $rank(\mathbf{A}) = 0$
- If $\mathbf{A}$ is $n \times n$, then $rank(\mathbf{A}) = n$ iff $|\mathbf{A}| \neq 0$
- kth minor of a matrix is the determinant of the matrix obtained by deleting all but k rows and k columns

Theorem 2

The rank of a matrix equals the order of its largest nonzero minor.

Theorem 3

The rank of a matrix is equal to the rank of its transpose.

Rank of a Matrix (2)

- To find matrix rank in R, an additional package is needed: **Matrix**
- This package comes pre-installed in the system library of R
- Still, needs loading

```
library(Matrix)
```

- To calculate the rank of a matrix:

```
rankMatrix(E)
```

- Note that we are using the default options, otherwise a lot of tweaking is possible (read the associated help)

Main Results on Linear Systems

Main Results on Linear Systems

- Consider the linear system $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is $m \times n$
- Define the following *augmented* matrix containing $\mathbf{A}$ and the vector $\mathbf{b}$:

$$\mathbf{A_b} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{pmatrix}$$

- It is clear that $rank(\mathbf{A_b}) \geq rank(\mathbf{A})$
- Also, $rank(\mathbf{A_b}) \leq rank(\mathbf{A}) + 1$

Main Results on Linear Systems (2)

Theorem 4

$\mathbf{Ax} = \mathbf{b}$ *has a solution* $\Leftrightarrow rank(\mathbf{A_b}) = rank(\mathbf{A})$.

- If $\mathbf{A}$ is $n \times n$, then this theorem shows that if $rank(\mathbf{A}) = n$, the system has a solution

Theorem 5

Suppose that the system has solutions with $rank(\mathbf{A}) = rank(\mathbf{A_b}) = k$.

(a) *If* $k < m$, *then* $m - k$ *equations are superfluous, i.e. any solution that satisfies the* k *equations corresponding to linearly independent rows will also satisfy the remaining ones*

(b) *If* $k < n$ *(i.e. less the number of unknowns), then* $n - k$ *variables can be chosen freely (**"degrees of freedom"**), and the remaining* k *variables are determined uniquely by that choice.*

Eigenvalues

Eigenvalues

- In many applications of dynamic economics, you have to calculate $\mathbf{A}^n\mathbf{x}$, where $\mathbf{x}$ is a non-zero vector
- This is not easy
- Suppose however that there exists a *scalar* λ such that

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \quad (*)$$

- In such a case, $\mathbf{A}^n\mathbf{x} = \mathbf{A}^{n-1}(\mathbf{A}\mathbf{x}) = \mathbf{A}^{n-1}(\lambda\mathbf{x}) = \ldots = \lambda^n\mathbf{x}$
- A non-zero[1] vector x that solves $(*)$ is called an **eigenvector**
- The associated λ is called an eigenvalue
- Eigenvalues have many important applications; they will be of big importance for us in particular in differential/difference equations/time series analysis

[1] Trivial (zero) solutions are of no practical interest.

Eigenvalues (2)

- Eigenvalues are also called *characteristic roots*, respectively eigenvectors are known also as *characteristic vectors*
- Note that if $\mathbf{x}$ is an eigenvector associated with an eigenvalue λ, then $\alpha\mathbf{x}$, where $\alpha \neq 0$ is a scalar, is also an eigenvector
- Eigenvalues are found by solving the matrix equation

$$(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} = \mathbf{0}$$

- This is a homogeneous system which has a solution only if $|\mathbf{A} - \lambda\mathbf{I}_n| = 0$
- Therefore, we have to solve the following *characteristic equation* to find the eigenvalues of $\mathbf{A}$;

$$p(\lambda) = |\mathbf{A} - \lambda\mathbf{I}_n| = 0$$

Eigenvalues (3)

- The polynomial that results in this problem is called the *characteristic polynomial*
- It is an nth-degree polynomial, therefore it has n roots (real or complex)
- Some of the roots can be repeated
- After the eigenvalues are calculated, for each of them the corresponding eigenvectors can be calculated
- In R, eigenvalues and eigenvectors can be calculated as follows:[2]

```
A <- matrix(c(0,1/2,0,0,0,1/3,6,0,0), nrow = 3)
eigval <- eigen(A)$values
eigvect <- eigen(A)$vectors
```

[2] Sysdsaeter et. al., p. 22, Example 2.

Eigenvalues (4)

Theorem 6

If $\mathbf{A}$ *is* $n \times n$ *with eigenvalues* $\lambda_1, \lambda_2, \ldots, \lambda_n$, *then*

(a) $|\mathbf{A}| = \lambda_1 \lambda_2 \ldots \lambda_n$

(b) $\mathrm{tr}(\mathbf{A}) = a_{11} + a_{22} + \ldots + a_{nn} = \lambda_1 + \lambda_2 + \ldots + \lambda_n$

Diagonalization

Diagonalization

- Let $\mathbf{A}$ and $\mathbf{P}$ be $n \times n$, and also assume that $\mathbf{P}^{-1}$ exists
- Consider the matrix $\mathbf{P}^{-1}\mathbf{AP}$; it has the following characteristic polynomial:

$$|\mathbf{P}^{-1}\mathbf{AP} - \lambda\mathbf{I}|$$

- But since $\lambda\mathbf{I} = \mathbf{P}^{-1}\mathbf{P}\lambda\mathbf{I} = \mathbf{P}^{-1}\lambda\mathbf{IP}$, we can write this characteristic polynomial as:

$$|\mathbf{P}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{P}|$$

- Using the fact that the determinant of the matrix product equals the product of determinants, and also the fact that $|\mathbf{P}^{-1}| = 1/|\mathbf{P}|$, we get

$$|\mathbf{P}^{-1}||(\mathbf{A} - \lambda\mathbf{I})||\mathbf{P}| = |(\mathbf{A} - \lambda\mathbf{I})|$$

- In other words, $\mathbf{A}$ and $\mathbf{P}^{-1}\mathbf{AP}$ have the same characteristic polynomial

Diagonalization (2)

- $\mathbf{A}_{n\times n}$ is diagonalizable if there exists an invertible $\mathbf{P}_{n\times n}$ matrix and a diagonal matrix $\mathbf{D}$ such that

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$$

Theorem 7

$\mathbf{A}_{n\times n}$ is diagonalizable iff it has a set of linearly independent eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. In such a case

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

where the eigenvectors are the columns of $\mathbf{P}$, and the lambdas are the corresponding eigenvalues.

- A direct implication of the theorem is that $\mathbf{A} = \mathbf{P}\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)\mathbf{P}^{-1}$. Therefore, for $m \in \mathbb{R}$

$$\mathbf{A}^m = \mathbf{P}\operatorname{diag}(\lambda_1^m, \lambda_2^m, \ldots, \lambda_n^m)\mathbf{P}^{-1}$$

Diagonalization (3)

- $\mathbf{P}$ is called an *orthogonal matrix* if $\mathbf{P}' = \mathbf{P}^{-1}$, i.e. $\mathbf{P}'\mathbf{P} = \mathbf{I}$
- If $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are the n columns of $\mathbf{P}$, then $\mathbf{x}_1', \mathbf{x}_2', \ldots, \mathbf{x}_n'$ will be the n rows of $\mathbf{P}'$
- But then $\mathbf{P}'\mathbf{P} = \mathbf{I}$ reduces to n equations of the type $\mathbf{x}_i'\mathbf{x}_j = 1$ for $i = j$ and $\mathbf{x}_i'\mathbf{x}_j = 0$ for $i \neq j$
- Therefore, $\mathbf{P}$ is orthogonal iff $\mathbf{x}_2, \ldots, \mathbf{x}_n$ all have length 1 and are mutually orthogonal
- In economics, we often encounter symmetric matrices (e.g. covariance matrices)
- The following theorem concerns such symmetric matrices

Diagonalization (4)

Theorem 8 (Spectral Theorem for Symmetric Matrices)

If $\mathbf{A}_{n\times n}$ is a symmetric matrix then

(a) $\lambda_1, \lambda_2, \ldots, \lambda_n$ *are all real*

(b) *Eigenvectors corresponding to different eigenvalues are orthogonal*

(c) *There exists an orthogonal matrix* $\mathbf{P}$ *such that*

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

The columns of $\mathbf{P}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ *are eigenvectors of length 1 corresponding to the* n *eigenvalues* $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Diagonalization (5)

- In R

```
A <- matrix(1:9, nrow = 3)
P <- eigen(A)$vectors
LAMBDA <- diag(eigen(A)$values)

P %*% LAMBDA %*% solve(P)
```

Differentiation

Gradient

- Take the function

$$y = f(x_1, x_2, \ldots, x_n)$$

- This function takes as an argument a vector $\mathbf{x}$ and is scalar-valued
- The vector of its partial derivatives with respect to each of the x's is called the *gradient (gradient vector)* and is written as

$$\nabla f(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial y}{\partial x_1} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{pmatrix}$$

Derivative of a Linear Function

- Take the vector of constants $\mathbf{a} = (a_1,\ a_2, \ldots, a_n)'$
- A linear function of the $\mathbf{x}$ vector can be written as

$$y = f(\mathbf{x}) = \mathbf{a}'\mathbf{x} = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n$$

- Take the vector of partial derivatives:

$$\nabla f(\mathbf{x}) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \mathbf{a}$$

Derivative of a Linear Function (2)

- Now take a set of linear functions $y_j = \mathbf{a}_j'\mathbf{x}, \quad j = 1, 2, \ldots, m$
- We can write all of them compactly as

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

where $\mathbf{a}_j'$ are the rows of the matrix $\mathbf{A}$

- Taking the gradient of each function leads to the following

$$\nabla y_j = \frac{\partial y_j}{\partial \mathbf{x}} = \mathbf{a}_j,$$

i.e. the transpose of the corresponding jth row of $\mathbf{A}$

Derivative of a Linear Function (3)

- We can also write

$$\begin{pmatrix} \partial y_1 / \partial \mathbf{x}' \\ \partial y_2 / \partial \mathbf{x}' \\ \vdots \\ \partial y_m / \partial \mathbf{x}' \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \vdots \\ \mathbf{a}_m' \end{pmatrix}$$

- Using the above,

$$\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}'} = \mathbf{A}; \quad \frac{\partial \mathbf{x}'\mathbf{A}'}{\partial \mathbf{x}} = \mathbf{A}'$$

Derivatives and Quadratic Forms

- Wikipedia:

  "...a quadratic form is a polynomial with terms all of degree two..."

- A quadratic form in $\mathbf{x}$ would be

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \sum_{i=1}^{n}\sum_{i=1}^{n} a_{ij}x_i x_j$$

  where $\mathbf{A}$ is $(n \times n)$

- Differentiate with respect to $\mathbf{x}$:

$$\frac{\partial \mathbf{x}'\mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = (\mathbf{A} + \mathbf{A}')\mathbf{x}$$

Complex Numbers

Definitions

- Suppose that the equation:

$$x^2 = -1$$

  has a solution
- Call this solution i, i. e. $i^2 = -1$
- That number i is know also as the **imaginary unit**.

Definition 2

A complex number is a number which can be expressed as $a + bi$, where $a, b \in \mathbb{R}^1$. a is the real part of the complex number, and b is its imaginary part.

- Let z be a complex number such that $z = a + bi$, then the following notation is used:

$$Re(z) = a$$
$$Im(z) = b$$

Operations with complex numbers

Let $z_1 = a + bi$, $z_2 = c + di$. The following operations can be carried out:

1. Addition: $z_1 + z_2 = (a + c) + (b + d)i$
2. Subtraction: $z_1 - z_2 = (a - c) + (b - d)i$
3. Multiplication: $(a + bi)(c + di) = (ac - bd) + (bc + ad)i$
4. Division: $\dfrac{a + bi}{c + di} = \dfrac{ac + bd}{c^2 + d^2} + \dfrac{bc - ad}{c^2 + d^2}i$

Operations with complex numbers (2)

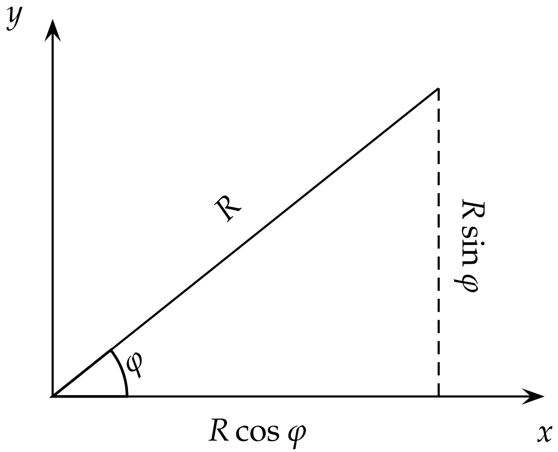
- Complex conjugate:

$$\bar{z} = a - bi$$

- (**Exercise:** find the complex conjugates of $z_1 + z_2, z_1 - z_2, z_1 z_2$ and z_1 / z_2.)
- Square root of a complex number:

$$\sqrt{z} = \sqrt{\frac{a + \sqrt{a^2 + b^2}}{2}} + sgn(b)\sqrt{\frac{-a + \sqrt{a^2 + b^2}}{2}}$$

Polar coordinates of z



Modulus and coordinates switching

- Modulus of a complex number:

$$R = |z| = \sqrt{a^2 + b^2}$$

- Cartesian coordinates from polar ones:

$$a = R\cos\varphi,\ b = R\sin\varphi$$

...and vice versa:

$$\varphi = \begin{cases} \arctan(b/a), & a > 0 \\ \arctan(b/a) + \pi, & a < 0,\ b \geq 0 \\ \arctan(b/a) - \pi, & a < 0,\ b < 0 \\ \pi/2, & a = 0,\ b > 0 \\ -\pi/2, & a = 0,\ b < 0 \\ \text{not defined}, & a = 0,\ b = 0 \end{cases}$$

- Using all this, we can write:

$$z = R(\cos\varphi + i\sin\varphi)$$

Euler's formula

- Stated:

$$z = e^{i\varphi}$$

- How it is derived?
- First, recall the idea of Taylor series expansion
- Let the function $f(x)$ be infinitely differentiable at the point a[3]
- Then $f(x)$ is equal to:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \dots$$

- MacLaurin series expansion is Taylor series expansion at the point 0:

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \dots$$

[3] a can be a real or a complex number

Euler's formula (2)

- Take the MacLaurin expansions of $\sin x$, $\cos x$, and e^z

$$
\begin{aligned}
\sin x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots \\
\cos x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots \\
e^z &= \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots
\end{aligned}
$$

Euler's formula (3)

- Now take $z = ix$ and substitute it in the MacLaurin expansion of e^z:

$$\begin{aligned} e^{ix} &= \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} = 1 + ix + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \cdots \\ &= 1 + ix - \frac{x^2}{2!} - i\frac{x^3}{3!} + \frac{x^4}{4!} + i\frac{x^5}{5!} - \cdots \\ &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots\right) \\ &= \cos x + i \sin x \end{aligned}$$

- Now it is easy to switch from x to φ in the formula

Euler's formula (4)

- $e^{i\varphi}$ can be considered as the unit complex number (i.e. having unit modulus); obviously, it lies on the unit circumference
- Any complex number $z = a + bi$ and its conjugate $\bar{z} = a - bi$ can then be written as:

$$z = R(\cos\varphi + i\sin\varphi) = Re^{i\varphi}$$
$$\bar{z} = R(\cos\varphi - i\sin\varphi) = Re^{-i\varphi}$$

This is especially useful in multiplying and dividing complex numbers:

$$z_1 z_2 = R_1 R_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2))$$
$$\frac{z_1}{z_2} = \frac{R_1}{R_2}(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2))$$

Euler's formula (5)

- de Moivre's theorem:

$$(\cos\varphi + i\sin\varphi)^n = \cos(n\varphi) + i\sin(n\varphi)$$

- Quick derivation of the result of the theorem:

$$(\cos\varphi + i\sin\varphi)^n = (e^{i\varphi})^n = e^{i(n\varphi)} = \cos(n\varphi) + i\sin(n\varphi)$$

- This is very convenient in compounding complex numbers in particular:

$$z^n = R^n e^{ni\varphi} = R^n(\cos(n\varphi) + i\sin(n\varphi))$$

Polynomial equations

- The degree of the polynomial defines the number of solutions (roots)
- For example, a 5th-degree polynomial has 5 solutions (roots)
- If a complex number is a root of a polynomial, then its conjugate is also a root
- How to solve in Matlab or GNU Octave:

```
p = [1 -2 8]
r = roots(p)
```

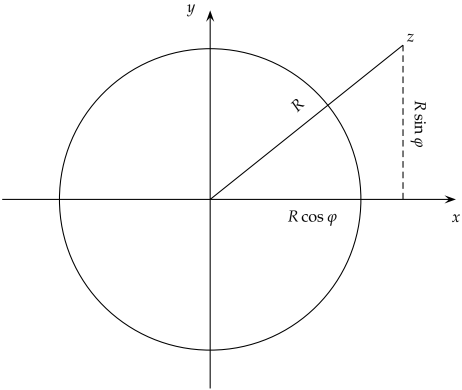
- How to solve the same in R:

```
polyroot(c(8, -2, 1))
```

- Note: In Maltab coefficients are specified in the decreasing order of powers; in R the increasing order is used

The unit circle

- It is important where a complex number lies with respect to the unit circle – inside, outside, or on the circumference



The above problem solved by hand

Solve the equation:

$$x^2 - 2x + 8 = 0$$

Solution:

The discriminant is $D = 4 - 4 \times 8 = -28 < 0$. Its square root is $\sqrt{D} = 2\sqrt{7 \times (-1)}$. This is equivalent to $2i\sqrt{7}$. The solutions of the equation are the two complex conjugates:

$$x_{1,2} = 1 \pm i\sqrt{7}$$

In polar form:

The modulus equals $R = \sqrt{1^2 + (\sqrt{7})^2} = \sqrt{8} = 2\sqrt{2}$. Then $\cos\varphi = a/R = \dfrac{1}{2\sqrt{2}} = \dfrac{\sqrt{2}}{2}$, and $\sin\varphi = b/R = \dfrac{\sqrt{7}}{2\sqrt{2}} = \dfrac{\sqrt{14}}{2}$.

Two more problems to solve

- Find $\ln(3 - 1.45i)$
- Find all solutions to $z^5 = 6i$
- Find all roots to $z^5 = -32$