Introduction to Matrix Algebra

Anton Velinov

Takeaways

- Vector spans, basis, dot products, orthogonality
- Basic matrix operations: transpose, trace, determinant, rank, inverse.
- Eigenvalues and Eigenvectors
- Matrix decompositions
- Special types of matrices: orthogonal, idempotent
- Definite matrices
- Kronecker product and matrix vectorization

Introduction I

Definition

An $(m \times n)$ *matrix*, $\mathbf{A}$ consists of an array of numbers ordered into m rows and n columns.

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

Or in shorthand $\mathbf{A} = [a_{ij}], i = 1 \ldots, m \quad j = 1, \ldots, n$.

Introduction II

Basics

- When m or n is 1, the matrix is called a *row* or a *column vector*, denoted as $\mathbf{a}'$ or $\mathbf{a}$ respectively
- A (1×1) matrix is a *scalar*
- An $(m \times m)$ matrix is a *square matrix*.

Note

Vectors and matrices are usually denoted in bold.

Vector Interpretation

- A vector is often interpreted as having a *magnitude* and a *direction* and is therefore, sometimes written as $\vec{v}$
- Usually, a vector starts from the point of origin, $\mathbf{0}$
- In vector addition, a vector is added to the tip of another vector
- In vector subtraction, a vector combines the tips of both vectors that are subtracted.

Example

- Draw the vector $\mathbf{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$
- Draw $\mathbf{a} + \mathbf{b}$
- Draw $\mathbf{a} - \mathbf{b}$.

Linear Dependence and Span

Definition: Linear Dependence

A set of vectors, $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ is said to be *linearly dependent* iff $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_n\mathbf{v_n} = 0$ for some $c_i s \in \mathbb{R}$, not all of which are zero (at least one c_i is non-zero).

Definition: Span

The *span* of the set of vectors, $S = \{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ is the (sub)space generated by every linear combination of the vectors in S. Formally, span(S)$= \{\sum_{i=1}^{n} c_i\mathbf{v_i}, c_i \in \mathbb{R}\}$.

Class Exercise

1. What is the span of $\mathbf{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$?
2. What is the span of $\mathbf{a}, \mathbf{b}$ and $\mathbf{c} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$?

Basis

Definition: Basis

A *basis* is the minimum set of vectors that spans a (sub)space. Note, this set is not unique and its vectors should be linearly independent.

A commonly used set of basis vectors for $\mathbb{R}^2$ is

$$B_{\mathbb{R}^2} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \equiv \{i, j\}$$

and for $\mathbb{R}^3$ is $B_{\mathbb{R}^3} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \equiv \{i, j, k\}$.

Class Exercise

Show that any vector in a (sub)space can be represented by a *unique* combination of the vectors in its basis. **Hint**: use proof by contradiction.

Dot Product and Orthogonality

Definition: Dot Product

A vector *dot product* or *inner product* between two *n*-dimentional vectors $\mathbf{a}$ and $\mathbf{b}$ is defined as $\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{n} a_i b_i$.

Note, the $L-2$ norm can be expressed as the square root of the dot product as $\|\mathbf{a}\|_2 = \sqrt{\mathbf{a} \cdot \mathbf{a}}$.

Angle between Vectors

The angle θ between two vectors, $\mathbf{a}$ and $\mathbf{b}$ is defined in terms of the dot product and vector lengths as

$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\|_2 \|\mathbf{b}\|_2 \cos\theta.$$

Definition: Orthogonality

Two (nonzero) vectors $\mathbf{a}$ and $\mathbf{b}$ are said to *orthogonal* if $\mathbf{a} \cdot \mathbf{b} = 0$ hence, the angle between them is 90°.

Summation

Definition

Matrix *addition* and *subtraction* is carried out for each corresponding element. I.e.

$$\mathbf{A} \pm \mathbf{B} = \mathbf{C}$$

means that

$$a_{ij} \pm b_{ij} = c_{ij} \quad \forall\, i, j.$$

Note

Only matrices with the *same* dimensions can be added (subtracted).

Multiplication

Definition

Let $\mathbf{A}$ be an $(n \times m)$ matrix and $\mathbf{B}$ an $(m \times k)$ matrix, then their *product* $\mathbf{AB} = \mathbf{C}$ is an $(n \times k)$ matrix with row i column j element given as

$$c_{ij} = \sum_{q=1}^{m} a_{iq} b_{qj}, \quad \forall\ i = 1, \ldots, n, j = 1, \ldots, k$$

Unlike matrix addition, matrix multiplication is not commutative

$$\mathbf{AB} \neq \mathbf{BA}.$$

Note

A matrix multiplied by a scalar, α affects each of its elements, $\alpha\mathbf{A} = \mathbf{C}$, where $\mathbf{C} = [\alpha a_{ij}]$.

Identity Matrix

Definition

An $(n \times n)$ *identity matrix*, denoted $\mathbf{I}_n$ is a diagonal matrix with ones on the diagonal and 0s elsewhere:

$$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Or equivalently $\mathbf{I}_n = \text{diag}(1, 1, \ldots, 1)$.

For any $(n \times m)$ matrix $\mathbf{A}$, it follows that $\mathbf{I}_n\mathbf{A} = \mathbf{A}$ and that $\mathbf{A}\mathbf{I}_m = \mathbf{A}$.

Matrix Transpose

Definition

The *transpose* of an $(n \times m)$ matrix, $\mathbf{A}$ denoted by $\mathbf{A}'$ or $\mathbf{A}^T$, is given as

$$\mathbf{A} = [a_{ij}], \quad \mathbf{A}' = [a_{ji}].$$

Properties

Transposition obeys the following properties

$$(\mathbf{A}')' = \mathbf{A}$$

$$(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$$

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'.$$

Symmetric Matrices

Definition

A matrix, $\mathbf{A}$ is said to be *symmetric* if $\mathbf{A}' = \mathbf{A}$.

Note

The product of two symmetric matrices is not necessarily symmetric.

Trace of a Matrix

Definition

The *trace* of an $(n \times n)$ matrix, $\mathbf{A}$ is the sum of its diagonal elements, i.e.

$$tr(\mathbf{A}) = \sum_{1=1}^{n} a_{ii}.$$

Note

The trace operator is *only* applicable to square matrices or the square matrix product of matrices.

Class Exercise

Show that for the $(m \times n)$ matrix, $\mathbf{A}$ and the $(n \times m)$ matrix, $\mathbf{B}$,

$$tr(\mathbf{AB}) = tr(\mathbf{BA}).$$

Trace Properties

1. Cyclic permutations are allowed: $tr(\mathbf{ABC}) = tr(\mathbf{BCA}) = tr(\mathbf{CAB})$ provided such products exist.
2. **However**, arbitrary permutations are not allowed, in general $tr(\mathbf{ABC}) \neq tr(\mathbf{ACB})$.
3. $tr(\mathbf{A} + \mathbf{B}) = tr(\mathbf{A}) + tr(\mathbf{B})$.

Class Exercise

For the scalar, γ, what is

$$tr(\gamma \mathbf{A})?$$

Determinant of a Matrix

Definition

The *determinant* of an $(n \times n)$ matrix, $\mathbf{A}$, denoted $|\mathbf{A}|$ is given as:

$$|\mathbf{A}| = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} |\mathbf{A}_{1j}|,$$

where $\mathbf{A}_{1j}$ is the $((n-1) \times (n-1))$ matrix formed by deleting row 1 column j from A

Note

The determinant is *only* applicable to square matrices.

Example

Using the above definition calculate the determinant of a (2×2) matrix.

Determinant Properties

1. $|\alpha \mathbf{A}| = \alpha^n |\mathbf{A}|$ for any scalar α and where n is the dimension of A
2. The determinant of a triangular matrix is simply the product of its diagonal elements
3. If a row/column of a matrix is multiplied by a scalar and added to another row/column, the new matrix will have the same determinant
4. If two rows/columns of a matrix are switched, the determinant changes signs
5. For the $(n \times n)$ matrices $\mathbf{A}$ and $\mathbf{B}$

$$|\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}|$$

6. $|\mathbf{A}'| = |\mathbf{A}|$.

Class Exercise

Determine $|\mathbf{I}_n|$ and verify properties 3 and 4 by means of a simple example.

Rank of a Matrix

Definition 1

The *rank* of a matrix is the number of linearly independent row or column vectors of the matrix.

Definition 2

The *rank* of an $(n \times m)$ matrix, $\mathbf{A}$, is the largest order of submatrices of $\mathbf{A}$ with a nonzero determinant, i.e. $rk(\mathbf{A}) \leq \min \{n, m\}$. If a square matrix is of full rank it is said to be nonsingular.

Rank Properties

1. Only for the null matrix, $\mathbf{0}$, $rk(\mathbf{0}) = 0$.
2. For $\mathbf{A}$ of dimension $(n \times m)$
   1. If $\mathbf{B}$ is an $(m \times l)$ matrix then $rk(\mathbf{AB}) \leq \min\{rk(\mathbf{A}), rk(\mathbf{B})\}$
   2. If $\mathbf{B}$ is an $(m \times l)$ matrix of rank m, then $rk(\mathbf{AB}) = rk(\mathbf{A})$
   3. If $\mathbf{C}$ is an $(k \times n)$ matrix of rank n, then $rk(\mathbf{CA}) = rk(\mathbf{A})$
3. For $\mathbf{A}$ and $\mathbf{B}$ of the same dimension

$$rk(\mathbf{A} + \mathbf{B}) \leq rk(\mathbf{A}) + rk(\mathbf{B}).$$

Class Exercise

Can you determine the rank of the $(n \times n)$ matrix $\mathbf{C}$ formed by the two nonzero $(n \times 1)$ column vectors $\mathbf{u}, \mathbf{v}$, $\mathbf{C} = \mathbf{uv}'$?

Inverse of a Matrix

Definition

The *inverse* of a nonsingular $(n \times n)$ matrix, $\mathbf{A}$, denoted as $\mathbf{A}^{-1}$ is given as

$$\mathbf{A}^{-1} = [(-1)^{i+j}|\mathbf{A}_{ji}|]/|\mathbf{A}|, \quad i, j = 1, \ldots, n$$

Example

Using the above definition calculate the inverse of a (2×2) matrix.

Inverse Properties

Properties of the inverse include

1. $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n$
2. $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$
3. $|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1}$
4. $(\gamma\mathbf{A})^{-1} = \mathbf{A}^{-1}/\gamma$ for any nonzero scalar
5. $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ where B is a nonsingular matrix with the same dimension as A.

Class Exercise

Verify $|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1}$ by means of a simple example.

Generalized Inverse of a Matrix

Definition

For a non-square (and possibly even singular) matrix, $\mathbf{A}$ of dimension $(n \times m)$ there exists a *generalized* or Moore-Penrose pseudoinverse, $\mathbf{A}^+$ such that

$$\mathbf{A}^+\mathbf{A} = \mathbf{I}.$$

For instance,

$$\mathbf{A}^+ = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'.$$

Eigenvalues and Eigenvectors I

Definition

The nonzero $(n \times 1)$ *Eigenvector*, $\mathbf{x}$ associated with the *Eigenvalue*, λ of an $(n \times n)$ matrix, $\mathbf{A}$ is related as follows:

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

Class Exercise

From the above definition why does it follow that an Eigenvalue, λ must be a number s.t.

$$|\mathbf{A} - \lambda\mathbf{I}_n| = 0?$$

Eigenvalues and Eigenvectors II

- An Eigenvector associated with an Eigenvalue is not unique and the following normalization is used to make it unique:

$$\mathbf{x}'_N \mathbf{x}_N = 1, \qquad \mathbf{x}_N = \frac{\mathbf{x}}{\|\mathbf{x}\|_2}$$

- The Eigenvalues of a triangular matrix are simply the values of its principle diagonal.

Question

Why is that so?

- If all $\lambda_i, i = 1, \ldots, n$ Eigenvalues of a matrix are distinct, then all the associated Eigenvectors, $\mathbf{x}_i, i = 1, \ldots, n$ are linearly independent

Eigenvalues and Eigenvectors III

Geometric Interpretation

Eigenvectors can be thought of as vectors that are invariant (up to scaling) to linear transformations. If all Eigenvectors are linearly independent they can serve as bases for coordinate systems.

Class Exercise

Find the eigenvalues and their associated eigenvectors of the following matrix

$$\mathbf{A} = \begin{bmatrix} 5 & 1 \\ 2 & 4 \end{bmatrix}.$$

Eigen Decomposition

If all of the Eigenvalues of an $(n \times n)$ matrix, $\mathbf{A}$ are distinct, it can be decomposed as follows:

- Collect all Eigenvalues in an $(n \times n)$ diagonal matrix, $\mathbf{\Lambda}$
- Collect all corresponding Eigenvectors to form the columns of an $(n \times n)$ matrix, $\mathbf{T}$
- From the Eigenvalue definition we have that $\mathbf{AT} = \mathbf{T\Lambda}$ and hence,

$$\mathbf{A} = \mathbf{T\Lambda T}^{-1}.$$

- This result can be used to derive the *powers* of a matrix,

$$\mathbf{A}^n = \mathbf{T\Lambda}^n\mathbf{T}^{-1}.$$

Class Exercise

Use this result to prove the statement:
"The determinant of a matrix with distinct Eigenvalues is equal to the product of its Eigenvalues."

Jordan Decomposition

Definition

If not all of the Eigenvalues of a matrix are distinct, we can still decompose the matrix by using the *Jordan decomposition*:

$$\mathbf{A} = \mathbf{MJM}^{-1}.$$

Where $\mathbf{J}$ is a matrix of the form $\mathbf{J} = \text{diag}(\mathbf{J}_1, \mathbf{J}_2, \ldots, \mathbf{J}_s)$, s stands for the number of linearly independent Eigenvectors and

$$\mathbf{J}_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix}.$$

Finally, $\mathbf{M}$ is an $(n \times n)$ nonsingular matrix.

Cholesky and LDL Decompositions

Cholesky decomposition

A (positive definite) symmetric matrix, **A** can be decomposed as

$$\mathbf{A} = \mathbf{L}\mathbf{L}'$$

where **L** is a lower triangular matrix. This is known as a *Cholesky decomposition* and it is unique if **A** is positive definite.

LDL decomposition

The same matrix can also be decomposed as

$$\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{L}'$$

where **D** is a diagonal matrix and **L** a lower triangular matrix. This is known as an *LDL decomposition* and is also unique if **A** is positive definite.

Orthogonal Matrices

Definition

An *orthogonal* matrix, $\mathbf{C}$ of dimension n has the property that

$$\mathbf{C}'\mathbf{C} = \mathbf{I}_n \quad \text{and} \quad \mathbf{C}\mathbf{C}' = \mathbf{I}_n.$$

If $\mathbf{A}$ is a symmetric matrix, then all of its Eigenvalues are real and their associated Eigenvectors are orthogonal. The Eigen decomposition of such a matrix can then be written as

$$\mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}'.$$

Idempotent Matrices I

Definition

An *idempotent* matrix, $\mathbf{A}$ has the property that

$$\mathbf{AA} = \mathbf{A}.$$

Class Exercise

Show that the symmetric idempotent matrix, $\mathbf{M}$ of dimension n has Eigenvalues all either 1 or 0.

- Hence, the Eigen decomposition of such a matrix is

$$\mathbf{M} = \mathbf{C\Lambda C}'$$

where $\mathbf{\Lambda} = \begin{bmatrix} \mathbf{I}_{n-k} & 0 \\ 0 & 0 \end{bmatrix}.$

Idempotent Matrices II

Class Exercise

Verify that for the symmetric idempotent matrix, $\mathbf{M}$ of rank $n-k$

$$rk(\mathbf{M}) = tr(\mathbf{M}).$$

Class Exercise

Verify that the following matrix is symmetric and idempotent

$$\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'.$$

Definite Matrices

Definition

A real symmetric matrix, $\mathbf{A}$ of dimension n is said to be *positive definite* if for any real n dimensional vector $\mathbf{v} \neq \mathbf{0}$ we have

$$\mathbf{v}'\mathbf{A}\mathbf{v} > 0.$$

Definition

An real symmetric matrix, $\mathbf{A}$ of dimension n is said to be *positive semidefinite* if for any real n dimensional vector $\mathbf{v}$ we have

$$\mathbf{v}'\mathbf{A}\mathbf{v} \geq 0.$$

Definition

Negative definite and negative semidefinite matrices are defined in an analogous manner.

Definite Matrices: Properties

1. All the Eigenvalues of a positive definite matrix are strictly positive
2. The inverse of a positive definite matrix is positive definite
3. Its determinant is positive and it is of full rank
4. For $\mathbf{A}$, an $(n \times n)$ positive definite matrix and $\mathbf{B}$, an $(n \times m)$ matrix with $rk(\mathbf{B}) = m$, $\mathbf{B}'\mathbf{A}\mathbf{B}$ is positive definite
5. Hence, $\mathbf{B}'\mathbf{B}$ is positive definite
6. The principle minors of a symmetric negative definite matrix alternate in sign

$$a_{11} < 0, a_{11}a_{22} - a_{12}^2 > 0, \ldots$$

Class Exercise

Prove properties 1 and 5.

Kronecker Product

Definition

The *Kronecker product* of the $(m \times n)$ matrix, $\mathbf{A}$ and the $(p \times q)$ matrix $\mathbf{B}$ is the $(mp \times nq)$ matrix $\mathbf{C}$:

$$\mathbf{C} = \mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2n}\mathbf{B} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}.$$

Kronecker Product Properties

1. In general $\mathbf{B} \otimes \mathbf{A} \neq \mathbf{A} \otimes \mathbf{B}$
2. $(\mathbf{A} \otimes \mathbf{B})' = \mathbf{A}' \otimes \mathbf{B}'$
3. $(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C})$
4. $(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = (\mathbf{A} \otimes \mathbf{C}) + (\mathbf{B} \otimes \mathbf{C})$
5. $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$
6. $(\mathbf{A} \otimes \mathbf{B})^{-1} = (\mathbf{A}^{-1} \otimes \mathbf{B}^{-1})$.

Matrix Vectorization

Definition

The *vectorization operator*, of the $(m \times n)$ matrix, $\mathbf{A}$ stacks all n columns of $\mathbf{A}$ in a single column vector, $\mathbf{a}$ of dimension mn:

$$vec(\mathbf{A}) = vec(\mathbf{a}_1 \quad \mathbf{a}_2 \quad \ldots \quad \mathbf{a}_n) = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{bmatrix} = \mathbf{a}.$$

Matrix Vectorization Properties

1. $vec(\mathbf{A} + \mathbf{B}) = vec(\mathbf{A}) + vec(\mathbf{B})$
2. $vec(\mathbf{ABC}) = [\mathbf{C}' \otimes \mathbf{A}]vec(\mathbf{B})$
3. $vec(\mathbf{ABC}) = (\mathbf{C}'\mathbf{B}' \otimes \mathbf{I})vec(\mathbf{A}) = (\mathbf{I} \otimes \mathbf{AB})vec(\mathbf{C})$
4. $vec(\mathbf{AB}) = vec(\mathbf{IAB}) = (\mathbf{B}' \otimes \mathbf{I})vec(\mathbf{A}) = (\mathbf{I} \otimes \mathbf{A})vec(\mathbf{B})$
5. $tr(\mathbf{AB}) = (vec(\mathbf{B}))'vec(\mathbf{A}')$

Example

Using these properties vectorize

$$\mathbf{Y} = \mathbf{X\Pi} + \mathbf{V}.$$

End of Theme 3

DIW GRADUATE CENTER

DIW Berlin – Deutsches Institut
für Wirtschaftsforschung e.V.
Mohrenstraße 58, 10117 Berlin
www.diw.de