Basics of Linear Algebra

Tim Cootes

Vectors
- An ordered list of \( n \) numbers
- Dimensionality: \( n \)

\[
x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}
\]

Transpose:

\[
x^T = (x_1 \ x_2 \ \ldots \ x_n)
\]

Geometric Interpretation
- Vector is difference between points in an \( n \)-dimensional space
- For instance, in 2D:

\[
\begin{pmatrix} x_1 \\ y_1 \\ \vdots \\ x_n \end{pmatrix}
\]

Manipulating Vectors
- Basic arithmetic operations act on each element independently:

\[
a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}
\]

\[
b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}
\]

\[
sa = \begin{pmatrix} sa_1 \\ sa_2 \\ \vdots \\ sa_n \end{pmatrix}
\]

\[
a + b = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix}
\]

Magnitude of a Vector
- Length of vector, written \(|a|\)
- By Pythagorus:

\[
|a| = \sqrt{\sum a_i^2}
\]

- In 2D:

\[
|a| = \sqrt{a_1^2 + a_2^2}
\]
Inner (“dot”) Product

- Dot product of two vectors is the sum of the products of matching elements:
  \[ \mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n \]
- Vectors must have the same number of elements

Properties of the Inner product

- If \( \mathbf{u} \) is a unit vector (length, \( |\mathbf{u}| = 1 \)), then \( \mathbf{a} \cdot \mathbf{u} \) gives the length of the component of \( \mathbf{a} \) along the direction of \( \mathbf{u} \)
  \[ \mathbf{a} \cdot \mathbf{u} = |\mathbf{a}| \cos \theta = |\mathbf{u}| \cos \theta \]

Matrices

- An ordered grid of numbers
- \( m \times n \)
  - \( m \) rows
  - \( n \) columns

\[ A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \]
- Note: \( n \)-dimensional vector is an \( n \times 1 \) matrix

Matrix Transpose

- Transpose: Flip about leading diagonal
- If \( A \) is \( n \times m \), its transpose, \( A^T \), is \( m \times n \)

Special Matrices

- Diagonal matrix:
  \[ \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & d_n \end{pmatrix} \]
- Symmetric matrix (must be square):
  \[ A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \]
  \[ a_{ij} = a_{ji} \]
Matrix Arithmetic

\[
\mathbf{A} + \mathbf{B} = \begin{pmatrix}
\mathbf{a}_1 + \mathbf{b}_1 & \mathbf{a}_2 + \mathbf{b}_2 & \cdots & \mathbf{a}_n + \mathbf{b}_n \\
\mathbf{a}_1 + \mathbf{b}_1 & \mathbf{a}_2 + \mathbf{b}_2 & \cdots & \mathbf{a}_n + \mathbf{b}_n \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{a}_1 + \mathbf{b}_1 & \mathbf{a}_2 + \mathbf{b}_2 & \cdots & \mathbf{a}_n + \mathbf{b}_n \\
\end{pmatrix}
\]

\(\mathbf{A} + \mathbf{B}\) must have same size

Matrix Multiplication

\[
\mathbf{A} \mathbf{B} = \begin{pmatrix}
\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\
\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\
\end{pmatrix}
\begin{pmatrix}
\mathbf{b}_1 \\
\mathbf{b}_2 \\
\vdots \\
\mathbf{b}_n \\
\end{pmatrix}
\]

\(j\)-th element of \(\mathbf{c}\) given by sum of products of \(j\)-th row of \(\mathbf{A}\) with elements of \(\mathbf{B}\)

\[
c_j = \sum_{i=1}^{m} \mathbf{a}_{ij} \mathbf{b}_i
\]

Matrix Identity

- Multiplying by identity leaves vector unchanged
- Identity, \(\mathbf{I}\), is square with zero everywhere except the leading diagonal:

\[
\mathbf{I} = \begin{pmatrix}
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{pmatrix}
\]

Geometric Interpretation

- If matrix \(\mathbf{A}\) is square \((n \times n)\), then it can be thought of as applying an affine geometric transformation to \(n\)-D vectors
- \(\mathbf{A}\) rotates, scales and shears them
- \(\mathbf{A} \mathbf{b} = \mathbf{c}\)
  - \(\mathbf{c}\) is a transformed version of \(\mathbf{b}\)

2D Examples

- \(\mathbf{a} = \begin{pmatrix} x \\ 0 \end{pmatrix}\) applies scaling of \(x\)
- \(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\) applies rotation of \(90^\circ\)
- \(\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}\) applies rotation of \(\theta\)
- \(\begin{pmatrix} x \\ 0 \end{pmatrix}\) applies shear along \(x\)

2D Examples

- \(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix}\)
- \(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix}\)

\(\mathbf{x} + (\mathbf{a} \cdot \mathbf{x}) = \mathbf{b}\)
2D Examples

Geometric Representation
- Every 2 x 2 matrix can be decomposed into a combination of rotations, scaling and shears.
- Every n x n matrix can be decomposed into a combination of n-dimensional scalings, rotations and shears.

Determinant of a Matrix
- The determinant, |A|, is a measure of the relative volume change applied by the transformation represented by A.
- It is the volume of the unit square/cube/hyper-cube after transforming by A.

Determinants
For 2D matrices, |A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc
For diagonal matrices, |D| = \begin{vmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & d_n \end{vmatrix} = d_1 d_2 \ldots d_n

Matrix Products
- The j-th element of C is dot product of j-th row of A with j-th column of B.

Matrix Products
AB = C
\[ c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \]
Sizes: (n x m)(m x p) = (n x p)
N cols of A matches N rows of B

\[ ABC = (AB)C = A(BC) \]
Sizes: (n x m)(m x p)(p x q) = (n x q)
For square matrices, |AB| = |A||B|
Unit (“Orthonormal”) Matrices

• If \( U \) is a square matrix such that \(|U| = 1 \) and \( U^T U = I \) then \( U \) is a “Unit” or “Orthonormal” matrix
• Such matrices apply pure rotations
  – No scaling or shearing

2D Unit Matrices

• 2D unit matrices have the form
  \[
  U = \begin{pmatrix}
  \cos \theta & -\sin \theta \\
  \sin \theta & \cos \theta
  \end{pmatrix}
  \]
  (they are pure rotations)

  - Applies a rotation of \( \theta \)
  - \( U^T \) applies a rotation of \( -\theta \)

Eigenvector Decomposition

• Any square, symmetric matrix \( M \) can be decomposed into a combination of pure rotations and simple scalings:

\[
M = U D U^T
\]

- Diagonal matrix
- Pure rotation
- Inverse of rotation
- Pure scaling along each axis

Transformations

• Thus the mapping \( \mathbf{x} \rightarrow M \mathbf{x} \)
  – Rotates the vector
  – Scales the vector by \( d_i \) along dimension \( i \)
  – Rotates the result back again

Eigenvectors

• The columns of the rotation matrix are the “Eigenvectors” of \( M \)
• The elements of the diagonal matrix are the “eigenvalues” of \( M \)

Each eigenvector \( \mathbf{u}_i \) has a corresponding eigenvalue \( d_i \)

Note: There are numerical algorithms for computing the eigenvectors/values of matrices

Determinants of symmetric matrices

• If \( M \) is symmetric,
  \[
  |M| = |UDU^T| = |U||D||U^T| = 1|D|= d_1, d_2, …, d_n