Lecture 3: Matrices

1 Matrices

A motivation for bringing in matrices is better expressing linear relations between variables.

Example 1. *The coordinates of a point in two different Cartesian coordinate systems have linear relations. Suppose in the first coordinate system, the point $P = (x_1, x_2, x_3)$, and in the second one, the point $P = (u_1, u_2, u_3)$. The relations between them can be described by linear equations, such as*

$$\begin{cases} u_1 = 2x_1 + x_2 + 5x_3 \\ u_2 = x_1 + 3x_2 + 2x_3 \\ u_3 = x_1 + 2x_2 + x_3 \end{cases}$$

The reason why their relation is linear is that each unit vector in the second coordinate system can be decomposed into three vectors along the directions of the unit vectors in the first coordinate system, i.e. each unit vector in the second coordinate system has a unique coordinate in the first system. Since x_1, x_2, x_3 are just scalars in three directions of the first coordinate system, the coordinates of the unit vectors of the second system can also be expressed with x_1, x_2, x_3, a linear combination of these three scalars. Then the coordinate of the point in the second coordinate system is also a linear combination of the unit vectors in the system, so the coordinate of the point in the second coordinate system is a linear combination of the coordinate of the point in the first one.
Then we can actually use matrix product to express the linear system:

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 5 \\ 1 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$U = AX$$

Such a matrix $A = \begin{bmatrix} 2 & 1 & 5 \\ 1 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix}$ is called a tranformation matrix, and such operation AX is called matrix product.

1.1 Definition of matrix product

Entries in a matrix product AB are the dot products of the corresponding rows in the matrix A and the corresponding columns in the matrix B respectively. Therefore, in order to apply the matrix product to the matrices A and B in this way AB,

$$\text{the number of columns in A} = \text{the number of rows in B}$$

In other words, the width of A must equal the height of B.

1.2 Property of matrix product

1.2.1 Associative property

The matrix product satisfies the associative property.

$$(AB)X = A(BX)$$

Therefore, when there are multiple tranformation matrices applied on a vector such as ABX, it means the tranformation matrix B is applied to X first, then the tranformation matrix A is applied to their result.

1.2.2 Commutative property

The matrix product doesn't satisfy the commutative property.

$$AB \neq BA$$

Not only in some cases AB is valid while BA doesn't make sense, but also when both AB and BA are valid, their results can be different.

1.3 Identity Matrix

The identity matrix I is defined as the matrix where

$$X = IX$$

Deriving from the definition, we know that all identity matrices are square matrices, and identity matrices always have the elements of the main diagonal (the list of entries $A_{i,j}$ where $i = j$) as 1, and other elements outside of the main diagonal as 0.

$$I_{4\times4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Example 2. *Find the transformation matrix to make a plane rotated by* $90°$ *counterclockwise.*
Since the transformation matrix represents a rotation of a plane, i.e. transforming each 2D point on the plane into another 2D point, the size of the transformation matrix is 2×2*:*

$$R = \begin{bmatrix} r_1 & r_2 \\ r_3 & r_4 \end{bmatrix}$$

Then we can use some points and their transformed results to figure out the elements:

$$A = \begin{bmatrix} 1 \\ 0 \end{bmatrix} A' = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} B' = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

After calculation,

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

We can verify the result by multiply itself by 4 times, which in effect doesn't rotate the plane at all, hence the result should be the identity matrix.

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

2 Inverse Matrix

2.1 Definition of inverse matrix

The inverse matrix of A, usually denoted as A^{-1}, is defined as

$$A^{-1}A = I$$
$$AA^{-1} = I$$

Deriving from the definition, only square matrices have inverse matrix. The matrices which have its corresponding inverse matrix are said to be invertible. Invert matrices can help solve matrix equations:

$$AX = B$$
$$A^{-1}AX = A^{-1}B$$
$$X = A^{-1}B$$

2.2 How to invert a matrix

The formula of inverse matrices is:

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

$adj(A)$ is the adjugate matrix of A, or the classical adjoint matrix of A.

Steps to invert a matrix:
1. Calculate the minor M of A. The entry on the position (i, j) of the minor, called the (i, j)-minor of A and denoted $M_{i,j}$, is the determinant of the matrices that results from deleting the row i and column j of A.
2. Calculate the cofactor matrix C of A. The entry on the position (i, j) of the cofactor matrix is the result of multiplying $M_{i,j}$ by $(-1)^{i+j}$.
3. Calculate the adjugate matrix $adj(A)$ of A by transposing the cofactor matrix of A. Transposing means switching rows and columns correspondingly.
4. Calculate the inverse matrix of A by dividing the adjugate matrix of A by the determinant of A.