Matrix Multiplication / A Review

Sources:

1. `Linear Algebra : Theory, Intuition, Code` author: Mike X Cohen, publisher: sincXpress

2. `No bullshit guide to linear algebra` author: Ivan Savov

Matrix * Matrix Multiplication

Matrix multiplication like

$$\underbrace{\mathbf{C}}_{m\times k} = \underbrace{\mathbf{A}}_{m\times n} \cdot \underbrace{\mathbf{B}}_{n\times k}$$

is defined if matrix $\mathbf{A}$ is of type $m \times n$ and matrix $\mathbf{B}$ is of type $n \times k$. The resulting matrix $\mathbf{C}$ is of type $m \times k$.

An element $c_{(i,j)}$ of matrix $\mathbf{C}$ is computed as the *dot product* of the i'th row vector of $\mathbf{A}$ and the j'th column vector of $\mathbf{B}$.

$$c_{(i,j)} = \sum_{l=1}^{n} a_{(i,l)} \cdot b_{(l,j)}$$

Obviously matrix multiplication is not commutative. While the product

$$\underbrace{\mathbf{A}}_{m\times n} \cdot \underbrace{\mathbf{B}}_{n\times k}$$

is defined and the result is a matrix of type $m \times k$ the product

$$\underbrace{\mathbf{B}}_{n\times k} \cdot \underbrace{\mathbf{A}}_{m\times n}$$

is not defined if $k \neq m$.

However apart from the algebraic definition of matrix multiplication there are other ways to express matrix multiplication. Depending on application context these other expressions may provide additional insight.

Review of the matrix-vector product

But first the product of a matrix $\mathbf{A}$ by a columns vector $\mathbf{x}$ is reviewed.

There are two cases to be considered:

1. right multiplication: column vector $\mathbf{x}$ multiplies matrix $\mathbf{A}$ from the right
2. left multiplication: row vector $\mathbf{x}^T$ multiplies matrix $\mathbf{A}$ from the left

right multiplication

$$\underbrace{\mathbf{A}}_{m \times n} \cdot \underbrace{\mathbf{x}}_{n \times 1} = \underbrace{\mathbf{b}}_{m \times 1}$$

The result of this multiplication is a column vector $\mathbf{b}$ with m elements.

The `i-th` element of column vector $\mathbf{b}$ is computed from this equation:

$$b_{(i)} = \sum_{l=1}^{n} a_{(i,l)} \cdot x_{(l)}$$

The column vector $\mathbf{b}$ is the weighted addition of column vectors of matrix $\mathbf{A}$. Weighting factors are taken from column vector $\mathbf{x}$. The `l-th` column of the matrix is weighted by the `l-th` element of vector $\mathbf{x}$.

left multiplication

$$\underbrace{\mathbf{x}^T}_{1 \times m} \cdot \underbrace{\mathbf{A}}_{m \times n} = \underbrace{\mathbf{b}^T}_{1 \times n}$$

The result of this multiplication is a row vector $\mathbf{b}^T$ with n elements.

The `l-th` element of row vector $\mathbf{b}^T$ is computed from this equation:

$$b_{(l)} = \sum_{i=1}^{m} a_{(i,l)} \cdot x_{(i)}$$

The row vector $\mathbf{b}^T$ is the weighted addition of row vectors of matrix $\mathbf{A}$. Weighting factors are taken from the row vector $\mathbf{x}^T$. The `i-th` row of the matrix is weigthed by the `i-th` element of row vector $\mathbf{x}^T$.

Now the equation for right multiplication between matrix and vector is reviewed under another perspective.

$$\underbrace{\mathbf{A}}_{m \times n} \cdot \underbrace{\mathbf{x}}_{n \times 1} = \underbrace{\mathbf{b}}_{m \times 1}$$

The element-wise computation of the column vector $\mathbf{b}$ had been defined by this equation:

$$b_i = \sum_{j=1}^{n} a_{(i,j)} \cdot x_j$$

Now we introduce two notations for row and column vectors of matrix $\mathbf{A}$.

i'the row vector of $\mathbf{A}$

$\mathbf{a}_{(i,j:)}$; the operator `j:` shall be read as range of `j` ; $1 \le j \le n$

j'the column vector of $\mathbf{A}$

$\mathbf{a}_{(i:,j)}$; the operator `i:` shall be read as range of `i` ; $1 \le i \le m$

With this notations it is possible to express the result vector $\mathbf{b}$ as the weighted addition of all column vectors $\mathbf{a}_{i:,j}$ of matrix $\mathbf{A}$. The weights are just the elements of vector $\mathbf{x}$.

$$\mathbf{b} = \sum_{j=1}^{n} \underbrace{\mathbf{a}_{(i:,j)}}_{j'th\ column\ vector} \cdot x_j$$

The latter equation is named `column perspective` of the matrix-vector product.

On the other hand it follows from the element wise computation of the matrix-vector product that vector element b_j can be computed from the dot / inner product of the `i'th` row vector $\mathbf{a}_{(i,j:)}$ and the column vector $\mathbf{x}$.

$$b_i = \mathbf{a}_{(i,j:)} \cdot \mathbf{x} = \underbrace{\sum_{j=1}^{n} a_{(i,j)} \cdot x_j}_{element-wise\ definition}$$

Outer product of two vectors

Let $\mathbf{a}$ be column vector and $\mathbf{b}$ a row vector.

$\mathbf{a}$ shall have `m` rows and $\mathbf{b}$ shall have `k` columns. The $\mathbf{a}$ is a special case of a $m \times 1$ matrix and $\mathbf{b}$ a special case of a $1 \times k$ matrix.

The matrix product (outer product) of these vectors / matrices yields a $m \times k$ matrix $\mathbf{C}$.

$$\mathbf{C} = \mathbf{a} \cdot \mathbf{b}$$

$$c_{(i,j)} = a_{(i,1)} \cdot b_{(1,j)} = a_{(i)} \cdot b_{(j)}$$

An example shall demonstrate the principle.

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \\ 2 & 4 & 2 \end{bmatrix}$$

Matrix - Matrix multiplication / column and row perspective

Starting with the *element-perspective* of matrix-matrix multiplication $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$

$$c_{(i,j)} = \sum_{l=1}^{n} a_{(i,l)} \cdot b_{(l,j)}$$

two equivalent forms

1. column perspective

2. row perspective

are defined. As before matrices have these dimensions:

1. matrix $\mathbf{A}$ is $m \times n$

2. matrix $\mathbf{B}$ is $n \times k$

3. matrix $\mathbf{C}$ is $m \times k$.

From the *element-perspective* of matrix-matrix multiplication the `j'th` column vector $\mathbf{c}_{(i:,j)}$ of matrix $\mathbf{C}$ is found as the weighted addition of column vectors $\mathbf{a}_{(i:,l)}$ of matrix $\mathbf{A}$. The multiplicative weigths are elements $b_{(l,j)}$ of the `j'th` column vector $\mathbf{b}_{(l:,j)}$ of matrix $\mathbf{B}$:

$$\mathbf{c}_{(i:,j)} = \sum_{l=1}^{n} \mathbf{a}_{(i:,l)} \cdot b_{(l,j)}$$

This equation is named the *column-perspective* of matrix-matrix multiplication.

A similar approach is used to find the *row-perspective*. From *element-perspective* we see that the `i'th` row vector of $\mathbf{c}_{(i,j:)}$ of matrix $\mathbf{C}$ can be expressed like this:

$$\mathbf{c}_{(i,j:)} = \sum_{l=1}^{n} a_{(i,l)} \cdot \mathbf{b}_{(l,j:)}$$

The `i'th` row vector of $\mathbf{c}_{(i,j:)}$ is seen to be computed from the weighted addition of the row vectors $\mathbf{b}_{(l,j:)}$ of matrix $\mathbf{B}$. The multiplicative weights $a_{(i,l)}$ are the elements of the `i'th` row vector $\mathbf{a}_{(i,l:)}$ of matrix $\mathbf{A}$.

Summary

1. the *element-perspective* yields the elements of matrix $\mathbf{C}$ from the elements of matrices $\mathbf{A}$ and $\mathbf{B}$.
2. the *column-perspective* yields the colum vectors of matrix $\mathbf{C}$ from a weighted addition of the columns vectors of matrix $\mathbf{A}$
3. the *row-perspective* yields the row vectors of matrix $\mathbf{C}$ from a weighted addition of row vectors of matrix $\mathbf{B}$

Matrix - Matrix multiplication / layer perspective

Another way to describe matrix multiplication is called the `layer perspective` (terminology taken from: `Linear Algebra : Theory, Intuition, Code` author: Mike X Cohen, publisher: sincXpress)

Starting with the *element-perspective* of matrix-matrix multiplication $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$ where the elements of $\mathbf{C}$ are computed from this equation:

$$c_{(i,j)} = \sum_{l=1}^{n} a_{(i,l)} \cdot b_{(l,j)}$$

$\mathbf{A}$ is of type $m \times n$ and matrix $\mathbf{B}$ is of type $n \times k$ thus the resulting matrix $\mathbf{C}$ is of type $m \times k$.

It can be shown that matrix $\mathbf{C}$ can be written as the sum of `n` *partial* matrices $\mathbf{C}_l$ of type $m \times k$

To see this the equation for matrix elements $c_{(i,j)}$ is re-written like this:

$$c_{(i,j)} = \underbrace{a_{(i,l=1)} \cdot b_{(l=1,j)}}_{c(l=1)_{(i,j)}} + \underbrace{a_{(i,l=2)} \cdot b_{(l=2,j)}}_{c(l=2)_{(i,j)}} + \ldots + \underbrace{a_{(i,l=n)} \cdot b_{(l=n,j)}}_{c(l=n)_{(i,j)}}$$

In this equation element $c(l=1)_{(i,j)}$ is the matrix element of partial matrix $\mathbf{C}_1$ and $c(l=2)_{(i,j)}$ the element of partial matrix $\mathbf{C}_2$ up to element $c(l=n)_{(i,j)}$ which is element of $\mathbf{C}_n$. Then the resulting matrix $\mathbf{C}$ is just the addition of these partial matrices:

$$\mathbf{C} = \sum_{l=1}^{n} \mathbf{C}_l$$

Looking at the definition of partial matrix element $c(l=1)_{(i,j)}$ we see that the partial matrix $\mathbf{C}_1$ is computed as the matrix product of the `1 st` column vector of matrix $\mathbf{A}$ multiplied from the right by the `1 st` row vector of matrix $\mathbf{B}$.

For some value `l` in the range $1 \le l \le n$ the partial matrix $\mathbf{C}_l$ is the matrix product of the `l th` column vector of matrix $\mathbf{A}$ multiplied from the right by the `l th` row vector of matrix $\mathbf{B}$.

example of layer perspective

$$\mathbf{C} = \begin{bmatrix} 3 & 0 & 3 \\ 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 4 & 4 \\ 0 & 1 \\ 4 & 1 \end{bmatrix}$$

$$\mathbf{C}_{l=1} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 4 & 4 \end{bmatrix} = \begin{bmatrix} 12 & 12 \\ 4 & 4 \end{bmatrix}$$

$$\mathbf{C}_{l=2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{C}_{l=3} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 4 & 1 \end{bmatrix} = \begin{bmatrix} 12 & 3 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{C} = \mathbf{C}_{l=1} + \mathbf{C}_{l=2} + \mathbf{C}_{l=3} = \begin{bmatrix} 24 & 15 \\ 4 & 5 \end{bmatrix}$$

Examples

Matrix-Vector Product

$$\mathbf{A} = \begin{bmatrix} -5 & 5 & 5 \\ -3 & -7 & 15 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} -1 \\ 2 \\ -2 \end{bmatrix}$$

$$\mathbf{c} = \mathbf{A} \cdot \mathbf{x}$$

element-wise computation of $\mathbf{c}$

$$\mathbf{c} = \begin{bmatrix} -5 \cdot (-1) + 5 \cdot 2 + 5 \cdot (-2) \\ -3 \cdot (-1) - 7 \cdot 2 + 15 \cdot (-2) \end{bmatrix} = \begin{bmatrix} 5 \\ -41 \end{bmatrix}$$

column-wise computation

$$\mathbf{c} = -1 \cdot \begin{bmatrix} -5 \\ -3 \end{bmatrix} + 2 \cdot \begin{bmatrix} 5 \\ -7 \end{bmatrix} - 2 \cdot \begin{bmatrix} 5 \\ 15 \end{bmatrix} = \begin{bmatrix} -5 \\ -41 \end{bmatrix}$$

In [8]:
```python
import numpy as np

# for the element-perspective the computation
# of the matrix product C = A * B
# is done using Numpy

A = np.array([[-2, 1, 2], [3, -1, 3]])
B = np.array([[-2, 1, 2, -1], [3, -1, 3, 2], [2, 1, -1, -2]])
C = np.matmul(A, B)

print(f"C = A*B:\n{C}")
```

```
C = A*B:
[[ 11  -1  -3   0]
 [ -3   7   0 -11]]
```

Matrix-Matrix Product

$$\mathbf{A} = \begin{bmatrix} -2 & 1 & 2 \\ 3 & -1 & 3 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} -2 & 1 & 2 & -1 \\ 3 & -1 & 3 & 2 \\ 2 & 1 & -1 & -2 \end{bmatrix}$$

Matrix $\mathbf{C}$ as defined by the matrix product

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

shall be computed usin all 3 methods:

1. element perspective
2. column perspective
3. row perspective

element perspective

computed with `Numpy`. See code example above.

$$\mathbf{C} = \begin{bmatrix} 11 & -1 & -3 & 0 \\ -3 & 7 & 0 & -11 \end{bmatrix}$$

---

column perspective

The column vectors of matrix $\mathbf{C}$ are computed from this equation:

$$\mathbf{c}_{(i:,j)} = \sum_{l=1}^{n} \mathbf{a}_{(i:,l)} \cdot b_{(l,j)}$$

$$\mathbf{C} = \Bigg[\left(-2 \cdot \begin{bmatrix} -2 \\ 3 \end{bmatrix} + 3 \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 2 \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right) \left(1 \cdot \begin{bmatrix} -2 \\ 3 \end{bmatrix} - 1 \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 1 \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right) \left(2 \cdot \begin{bmatrix} -2 \\ 3 \end{bmatrix} + 3 \cdot \right.$$

$$\mathbf{C} = \begin{bmatrix} 11 & -1 & -3 & 0 \\ -3 & 7 & 0 & -11 \end{bmatrix}$$

---

row perspective

Row vectors of matrix $\mathbf{C}$ are computed from this equation:

$$\mathbf{c}_{(i,j:)} = \sum_{l=1}^{n} a_{(i,l)} \cdot \mathbf{b}_{(l,j:)}$$

$$\mathbf{C} = \begin{bmatrix} -2 \cdot \begin{bmatrix} -2 & 1 & 2 & -1 \end{bmatrix} + 1 \cdot \begin{bmatrix} 3 & -1 & 3 & 2 \end{bmatrix} + 2 \cdot \begin{bmatrix} 2 & 1 & -1 & -2 \end{bmatrix} \\ 3 \cdot \begin{bmatrix} -2 & 1 & 2 & -1 \end{bmatrix} - 1 \cdot \begin{bmatrix} 3 & -1 & 3 & 2 \end{bmatrix} + 3 \cdot \begin{bmatrix} 2 & 1 & -1 & -2 \end{bmatrix} \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 11 & -1 & -3 & 0 \\ -3 & 7 & 0 & -11 \end{bmatrix}$$

```
In [ ]:
```