# I.2 Matrix-Matrix Multiplication $AB$

$$\begin{bmatrix} 1 & 2 & 4 \\ 1 & 4 & 3 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 & 1 \\ 2 & 3 & 4 \\ 3 & 4 & 1 \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & c_{23} \\ \cdot & \cdot & \cdot \end{bmatrix}$$

$$A \qquad B \qquad = \qquad C$$

$$c_{23} = 1 \cdot 1 + 4 \cdot 4 + 3 \cdot 1 = 20$$

The <u>inner-product</u> formula for $C = AB$ is

$$c_{ij} = a_{i1} b_{1j} + \cdots + a_{in} b_{nj} = \sum_{k=1}^{n} a_{ik} b_{kj} .$$

(<u>ie</u> <u>row</u> of $A$ times <u>column</u> of $B$)

Another way to mult. $A$ and $B$ is using <u>outer-products</u> (<u>ie</u> <u>column</u> of $A$ times <u>row</u> of $B$).

<u>Example:</u> $u = \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix}$, $v = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}$

The <u>inner-product</u> of $u$ and $v$ is:

$$u^T v = \begin{bmatrix} 1 & 3 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} = 2 \cdot 1 + 1 \cdot 3 + 4 \cdot 3 = 17$$

The <u>outer-product</u> of $u$ and $v$ is:

$$u v^T = \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix} \begin{bmatrix} 2 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 4 \\ 6 & 3 & 12 \\ 6 & 3 & 12 \end{bmatrix} \leftarrow \text{rank} = 1$$

---

The <u>outer-product</u> formula for $AB$ is:

$$\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \begin{bmatrix} b_1^T \\ b_2^T \\ b_3^T \end{bmatrix} = a_1 b_1^T + a_2 b_2^T + a_3 b_3^T$$

( $a_1, a_2, a_3$ are the columns of $A$,
$b_1^T, b_2^T, b_3^T$ are the rows of $B$ )

$$\boxed{AB = \text{sum of rank one matrices}}$$

**Example:** Use both inner- and outer-product formulas to multiply the matrices.

$$\begin{bmatrix} 4 & 3 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 4\cdot4+3\cdot1 & 4\cdot0+3\cdot3 \\ 2\cdot4+0\cdot1 & 2\cdot0+0\cdot3 \end{bmatrix}$$

$$= \begin{bmatrix} 19 & 9 \\ 8 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 3 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}\begin{bmatrix} 4 & 0 \end{bmatrix} + \begin{bmatrix} 3 \\ 0 \end{bmatrix}\begin{bmatrix} 1 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 16 & 0 \\ 8 & 0 \end{bmatrix} + \begin{bmatrix} 3 & 9 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 19 & 9 \\ 8 & 0 \end{bmatrix}$$

---

## Matrix Factorizations

In the previous section we saw how to factor $A$ into $C$ times $R$.

$$A = CR = c_1 r_1^T + \cdots + c_k r_k^T$$

Factoring a matrix is the reverse of multiplying matrices and usually takes more effort and time.

Using the outer-product formula for matrix multiplication, the factorization splits the matrix up in a number of rank-one pieces.

We can use this to identify the most important pieces in our data, revealing structure and information in our data.

Six important factorizations are:

1. $A = LU$ (from Gaussian elimination)

   $L$ is lower-triangular with ones on the diagonal.

   $U$ is upper-triangular.

2. $A = LDL^T$ and $A = R^T R$ (Cholesky)

   If $A$ is symmetric, $A = LU$ becomes $A = LDL^T$, where $D$ is a diagonal matrix. If the eigenvalues of $A$ are all positive then $A = R^T R$ is possible.

③ $A = QR$ (from orthogonalizing the columns of $A$, as in "Gram-Schmidt")

$Q$ is orthogonal ($Q^TQ = I$).

$R$ is upper-triangular.

④ $A = Q\Lambda Q^T$ (eigenvalue decomposition)

$A$ is symmetric, so eigenvalues are real.

$Q$ is orthogonal, so $AQ = Q\Lambda$.

$\Lambda$ is diagonal. Thus, $Aq_i = \lambda_i q_i$,

so the columns of $Q$ are the eigenvectors and the diagonal entries of $\Lambda$ are the eigenvalues.

⑤ $A = X\Lambda X^{-1}$ (diagonalization)

If $A$ is nonsymmetric but has lin. indep. eigenvectors, then $A$ can be diagonalized.

⑥ $A = U\Sigma V^T$ (singular value decomposition)
$(m \times n)$

$U$ and $V$ are othogonal.

$\Sigma$ is diagonal with singular values

$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ on its diagonal.
($r = \text{rank}(A)$)

Thus, $Av_i = \sigma_i u_i$.

$u_i$ are the left-singular vectors.
$v_i$ are the right-singular vectors.

---

Big Idea: The outer-product multiplication formula tells us that

$$A = U\Sigma V^T = \sigma_1 u_1 v_1^T + \cdots + \sigma_r u_r v_r^T.$$

Since $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, we see that $\sigma_1 u_1 v_1^T$ is the most important rank-one piece of $A$.

---

(Numerical demonstration)