Chapter 6: Matrix Multiplication

Ketan Rajawat

Outline

Introduction

- Non-commutativity
- Properties
- Multiplication of Block Matrices

Interpretations

Other matrix operations

- Trace
- Hadamard Product

Complexity of Matrix multiplication

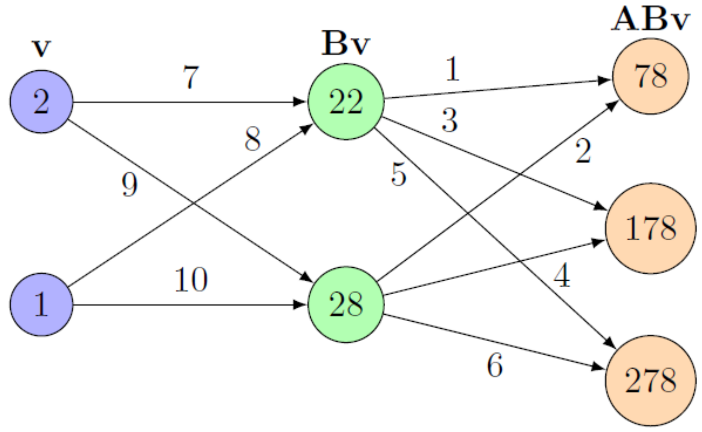
Introduction

$$\underset{m \times p}{\mathbf{C}} = \underset{m \times n}{\mathbf{A}}\ \underset{n \times p}{\mathbf{B}}$$

$$C_{ij} = \sum_{k=1}^{n} A_{ik} \cdot B_{kj}$$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 7 & 8 \\ 9 & 10 \end{bmatrix} = \begin{bmatrix} 25 & 28 \\ 57 & 64 \\ 89 & 100 \end{bmatrix}$$

Matrix-matrix-vector product



Special cases

- Consider two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$

- Inner product $\mathbf{u}^T\mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \langle \mathbf{u}, \mathbf{v} \rangle$

- Outer product $(\mathbf{u}\mathbf{v}^T)_{ij} = u_i v_j$
  - Note: $\mathbf{u}\mathbf{v}^T \neq \mathbf{v}\mathbf{u}^T$

- Multiplicative identity: $\mathbf{AI} = \mathbf{A} = \mathbf{IA}$

Non-commutativity

- Matrix multiplication is not commutative
- Consider $m \times n$ matrix **A** and $p \times q$ matrix **B**, then
  - $\mathbf{AB}$ only defined when $n = p$
  - $\mathbf{BA}$ only defined when $m = q$
  - $\mathbf{AB} \neq \mathbf{BA}$ even when $n = p$ and $m = q$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix} \qquad \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix}$$

Properties

- Associativity $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
- Associativity with scalar multiplication $(c\mathbf{A})\mathbf{B} = c(\mathbf{AB}) = \mathbf{A}(c\mathbf{B})$
- Distributivity with addition $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
- Implication: $(\mathbf{A} + \mathbf{B})(\mathbf{C} + \mathbf{D})$

$$= \mathbf{A}(\mathbf{C} + \mathbf{D}) + \mathbf{B}(\mathbf{C} + \mathbf{D})$$
$$= \mathbf{AC} + \mathbf{AD} + \mathbf{BC} + \mathbf{BD}$$

Transpose of Product

$$\underset{m\times p}{\mathbf{C}} = \underset{m\times n}{\mathbf{A}}\ \underset{n\times p}{\mathbf{B}}$$

- Claim: $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$

Transpose

$$[(\mathbf{AB})^T]_{ij} = [\mathbf{AB}]_{ji} = \sum_{k=1}^{n} A_{jk}B_{ki}$$

$$[\mathbf{B}^T\mathbf{A}^T]_{ij} = \sum_{k=1}^{n}[\mathbf{B}^T]_{ik}[\mathbf{A}^T]_{kj} = \sum_{k=1}^{n} B_{ki}A_{jk} = \sum_{k=1}^{n} A_{jk}B_{ki}$$

Transpose

For scalars $ab = ba$

Multiplication of Block Matrices

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}$$

$$\mathbf{AB} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{bmatrix}$$

- Sizes must be compatible
- number of columns in $\mathbf{A}_{11}$
  $=$ number of rows in $\mathbf{B}_{11}$
- …

Row/Column Interpretations

$$\mathbf{AB} = \mathbf{A}[\mathbf{b}_1 \quad \mathbf{b}_2 \quad \dots \quad \mathbf{b}_p] = [\mathbf{Ab}_1 \quad \mathbf{Ab}_2 \quad \dots \quad \mathbf{Ab}_p]$$

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_{1\cdot} \\ \mathbf{a}_{2\cdot} \\ \vdots \\ \mathbf{a}_{m\cdot} \end{bmatrix} \mathbf{B} = \begin{bmatrix} \mathbf{a}_{1\cdot}\mathbf{B} \\ \mathbf{a}_{2\cdot}\mathbf{B} \\ \vdots \\ \mathbf{a}_{m\cdot}\mathbf{B} \end{bmatrix}$$

Inner Product Interpretation

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_{1}. \\ \mathbf{a}_{2}. \\ \vdots \\ \mathbf{a}_{m}. \end{bmatrix}$$

$$[\mathbf{AB}]_{ij} = \mathbf{a}_{i}.\mathbf{b}_{j} = \left\langle\, \mathbf{a}_{i}^{T}.,\mathbf{b}_{j} \right\rangle$$

$$\mathbf{B} = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \dots \quad \mathbf{b}_p]$$

Gram Matrix

$$\mathbf{A}^T = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix}$$

$$\mathbf{G} = \mathbf{A}^T\mathbf{A}$$

$$(\mathbf{G})_{ij} = \mathbf{a}_i^T\mathbf{a}_j = \langle\, \mathbf{a}_i, \mathbf{a}_j \rangle$$

$$\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_n]$$

- Note: $\mathbf{G}$ is symmetric

$$\mathbf{G}^T = (\mathbf{A}^T\mathbf{A})^T = \mathbf{A}^T(\mathbf{A}^T)^T = \mathbf{A}^T\mathbf{A} = \mathbf{G}$$

Outer Product Interpretation

$$\mathbf{AB} = \mathbf{a}_1\mathbf{b}_{1\cdot} + \mathbf{a}_2\mathbf{b}_{2\cdot} + \cdots + \mathbf{a}_n\mathbf{b}_{n\cdot}$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{b}_{1\cdot} \\ \mathbf{b}_{2\cdot} \\ \vdots \\ \mathbf{b}_{n\cdot} \end{bmatrix}$$

Each summand is a matrix

$$\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \ldots \quad \mathbf{a}_n]$$

Matrix Operations: Trace

- Trace is the sum of diagonal entries of a square matrix

$$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} \mathbf{A}_{ii}$$

- Only defined for square matrices

$$\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$$

$$\mathrm{tr}(\alpha\mathbf{A}) = \alpha\mathrm{tr}(\mathbf{A})$$

Trace Property

- Claim: $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$

$$[\mathbf{AB}]_{ii} = \sum_{j=1}^{n} A_{ij}B_{ji}$$

$$\mathrm{tr}(\mathbf{AB}) = \sum_{i=1}^{m}[\mathbf{AB}]_{ii} = \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}B_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{m} B_{ji}A_{ij} = \sum_{j=1}^{n}[\mathbf{BA}]_{jj} = \mathrm{tr}(\mathbf{BA})$$

For scalars $ab = ba$

- Implication:

$$\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}\big(\mathbf{A}(\mathbf{BC})\big) = \mathrm{tr}(\mathbf{BCA}) = \mathrm{tr}\big(\mathbf{B}(\mathbf{CA})\big) = \mathrm{tr}(\mathbf{CAB})$$

Hadamard Product

- Aka: entry-wise or element-wise product

$$(\mathbf{A} \odot \mathbf{B})_{ij} = A_{ij} B_{ij}$$

- Key properties:
  - Commutativity $\mathbf{A} \odot \mathbf{B} = \mathbf{B} \odot \mathbf{A}$
  - Distributivity $\mathbf{A} \odot (\mathbf{B} + \mathbf{C}) = (\mathbf{A} \odot \mathbf{B}) + (\mathbf{A} \odot \mathbf{C})$
  - Scaling $k(\mathbf{A} \odot \mathbf{B}) = (k\mathbf{A}) \odot \mathbf{B} = \mathbf{A} \odot (k\mathbf{B})$

Complexity of matrix operations

Operation	**Flop count**
Trace	n
Hadamard Product	n^2
Matrix-Matrix Multiplication	$2n^3$

Matrix multiplication: each element requires n multiplications and $n-1$ additions, so total flops $=$ $n^2(2n-1) \approx 2n^3$

Thank You

Next: Inner Product