Linear Algebra Overview

Instructor: Xiaoqian Wang

8/23/2024

Slides prepared based on the Lectures slides of Linear Algebra from
https://www.deeplearningbook.org/lecture_slides.html

Scalars

- A scalar is a single number
- Integers, real numbers, rational numbers, etc.
- **We denote it with italic font:**

$$a, n, x$$

(Goodfellow 2016)

Vectors

- A vector is a 1-D array of numbers:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \tag{2.1}$$

- Can be real, binary, integer, etc.

- **We denote it with bold lowercase letter:**

$$\mathbf{x} \in \mathbb{R}^n$$

(Goodfellow 2016)

Matrices

- A matrix is a 2-D array of numbers:

$$\begin{array}{cc} & \text{Row} \\ & \\ \end{array}$$

$$\begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix}. \tag{2.2}$$

Row

Column

- **We denote it with bold uppercase letter:**

$$\boldsymbol{A} \in \mathbb{R}^{m \times n}$$

(Goodfellow 2016)

Tensors

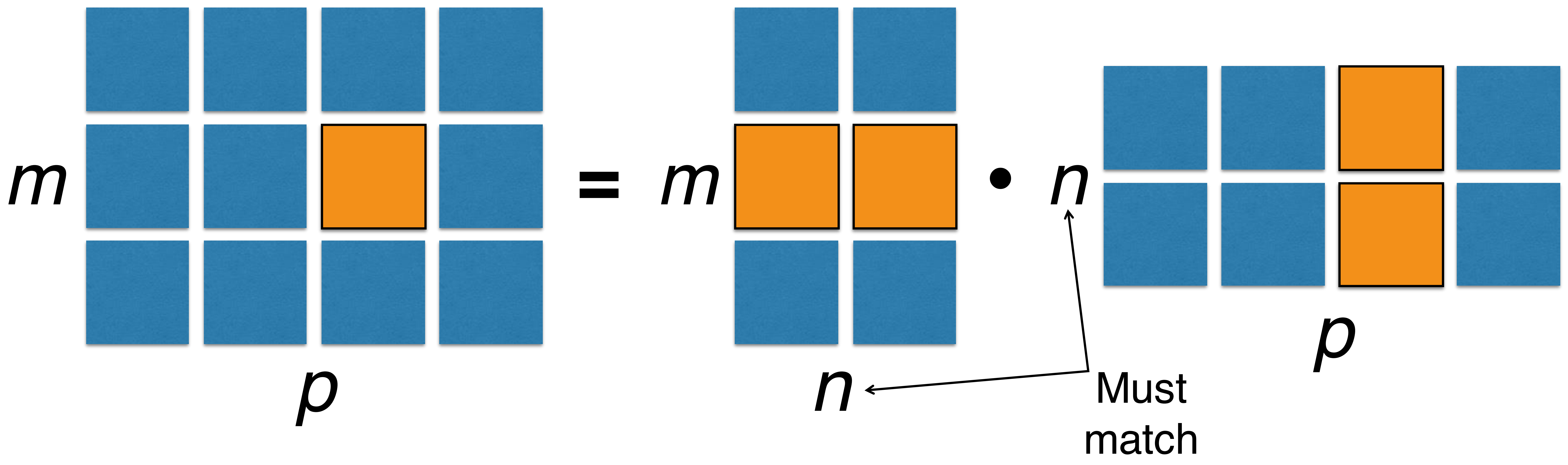
- A tensor is an array of numbers, that may have
  - zero dimensions, and be a scalar
  - one dimension, and be a vector
  - two dimensions, and be a matrix
  - or more dimensions.

(Goodfellow 2016)

Matrix (Dot) Product

$$\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}. \tag{2.4}$$

$$C_{i,j} = \sum_k A_{i,k} B_{k,j}. \tag{2.5}$$

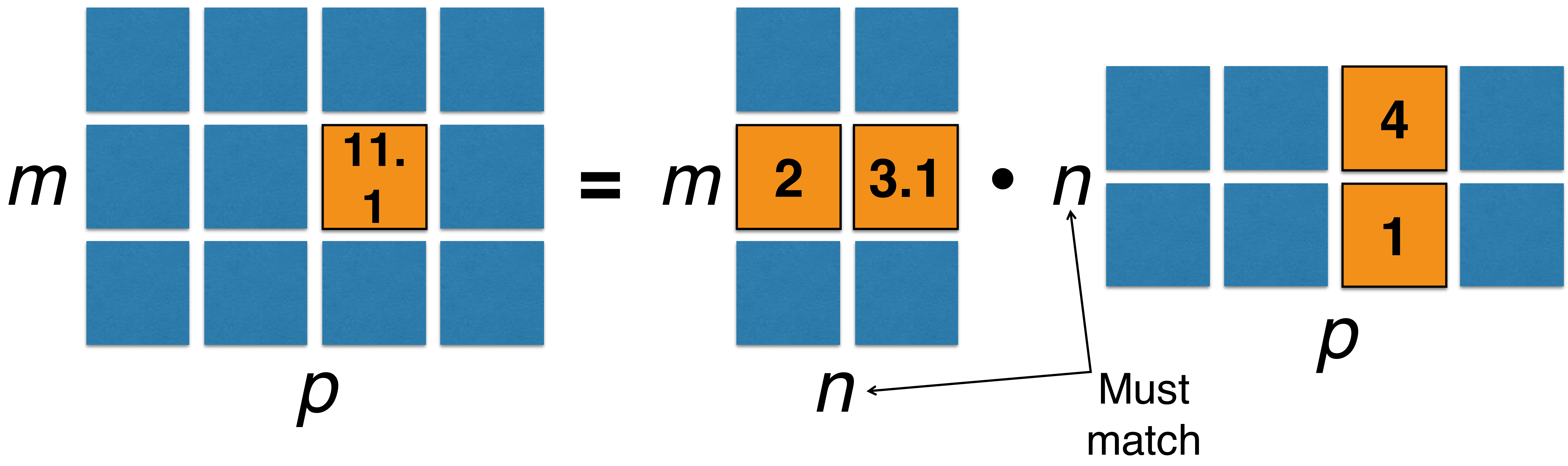


(Goodfellow 2016)

Matrix (Dot) Product Example

$$C = AB. \tag{2.4}$$

$$C_{i,j} = \sum_k A_{i,k} B_{k,j}. \tag{2.5}$$



Identity Matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Figure 2.2: *Example identity matrix*: This is $\boldsymbol{I}_3$.

$$\forall \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{I}_n \boldsymbol{x} = \boldsymbol{x}. \tag{2.20}$$

(Goodfellow 2016)

Matrix Transpose

$$(\boldsymbol{A}^\top)_{i,j} = A_{j,i}. \tag{2.3}$$

$$\boldsymbol{A} = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \\ A_{3,1} & A_{3,2} \end{bmatrix} \Rightarrow \boldsymbol{A}^\top = \begin{bmatrix} A_{1,1} & A_{2,1} & A_{3,1} \\ A_{1,2} & A_{2,2} & A_{3,2} \end{bmatrix}$$

Figure 2.1: The transpose of the matrix can be thought of as a mirror image across the main diagonal.

$$(\boldsymbol{AB})^\top = \boldsymbol{B}^\top \boldsymbol{A}^\top. \tag{2.9}$$

Special matrix: a matrix A is symmetric if A = A^T

(Goodfellow 2016)

Norms

- Functions that measure how “large” a vector is
- Similar to a distance between zero and the point represented by the vector
    - $f(\boldsymbol{x}) = 0 \Rightarrow \boldsymbol{x} = \boldsymbol{0}$
    - $f(\boldsymbol{x} + \boldsymbol{y}) \leq f(\boldsymbol{x}) + f(\boldsymbol{y})$ (the *triangle inequality*)
    - $\forall \alpha \in \mathbb{R}, f(\alpha \boldsymbol{x}) = |\alpha| f(\boldsymbol{x})$

(Goodfellow 2016)

Norms

- L^p norm

$$||\boldsymbol{x}||_p = \left(\sum_i |x_i|^p \right)^{\frac{1}{p}}$$

- Most popular norm: L2 norm, p=2 $\quad \| \mathbf{x} \|_2 = (\sum_i x_i^2)^{\frac{1}{2}}$
- L1 norm, p=1: $||\boldsymbol{x}||_1 = \sum_i |x_i|. \qquad (2.31)$
- Max norm, infinite p: $||\boldsymbol{x}||_\infty = \max_i |x_i|. \qquad (2.32)$
- Dot product of two vectors: $\boldsymbol{x}^\top \boldsymbol{y} = ||\boldsymbol{x}||_2 ||\boldsymbol{y}||_2 \cos\theta, \qquad (2.34)$

Special vector: a vector v is a unit vector if $||v||_2 = 1$

(Goodfellow 2016)

Norms

- Frobenius norm

$$\parallel \mathbf{A} \parallel_F = \left(\sum_i \sum_j A_{ij}^2\right)^{\frac{1}{2}}$$

Trace

$$\mathrm{Tr}(\boldsymbol{A}) = \sum_i \boldsymbol{A}_{i,i}. \tag{2.48}$$

$$\mathrm{Tr}(\boldsymbol{ABC}) = \mathrm{Tr}(\boldsymbol{CAB}) = \mathrm{Tr}(\boldsymbol{BCA}) \tag{2.51}$$

$$\mathrm{Tr}(\mathbf{A}) = \mathrm{Tr}(\mathbf{A}^{\mathrm{T}})$$

(Goodfellow 2016)

Matrix Inversion

- Matrix inverse:

$$A^{-1}A = I_n. \tag{2.21}$$

- **Special matrix:** a matrix A is an orthogonal matrix if $A^TA = I$, $AA^T = I$, $A^{-1} = A^T$

(Goodfellow 2016)

Invertibility

- Matrix can't be inverted if…
  - More rows than columns
  - More columns than rows
  - Redundant rows/columns ("linearly dependent", "low rank")

(Goodfellow 2016)

Eigendecomposition

- Eigenvector and eigenvalue for a <u>square</u> matrix A:

$$\boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}. \tag{2.39}$$

- Eigendecomposition of a diagonalizable matrix:

$$\boldsymbol{A} = \boldsymbol{V} \operatorname{diag}(\boldsymbol{\lambda}) \boldsymbol{V}^{-1}. \tag{2.40}$$

- Every real symmetric matrix has a real, orthogonal eigendecomposition:

$$\boldsymbol{A} = \boldsymbol{Q} \boldsymbol{\Lambda} \boldsymbol{Q}^{\top}, \tag{2.41}$$

(Goodfellow 2016)

Intuition of Eigendecomposition

- Understand eigendecomposition with visual examples
  - https://setosa.io/ev/eigenvectors-and-eigenvalues/

Singular Value Decomposition

- Similar to eigendecomposition
- More general; matrix need not be square

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^\top. \tag{2.43}$$

Computing the Pseudoinverse

The SVD allows the computation of the pseudoinverse:

$$\boldsymbol{A}^{+} = \boldsymbol{V}\boldsymbol{D}^{+}\boldsymbol{U}^{\top}, \tag{2.47}$$

Take reciprocal of non-zero entries

(Goodfellow 2016)

Learning linear algebra

- Do a lot of practice problems
- Start out with lots of summation signs and indexing into individual entries
- Eventually you will be able to mostly use matrix and vector product notation quickly and easily

(Goodfellow 2016)

Next Lecture

- Overview of Probability