Mathematics for AI

Lecture 3

Matrices

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 7 & 2 & -1 & 3 \\ 1.5 & 7 & 4 & 5 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} a[0,0] & a[0,1] & a[0,2] & a[0,3] \\ a[1,0] & a[1,1] & a[1,2] & a[1,3] \\ a[2,0] & a[2,1] & a[2,2] & a[2,3] \end{bmatrix}$$

$$\in \mathbb{R}^{3\times 4} \qquad \in \mathbb{Z}^{m\times n} \qquad \in \mathbb{C}^{m\times n}$$

Vectors as special matrices?

$$v = \begin{bmatrix}1\\0.5\\7\\2\end{bmatrix} \in \mathbb{R}^4 \quad \in \mathbb{R}^n \quad \in \mathbb{C}^n$$

column vector $\begin{bmatrix}1\\0.5\\7\\2\end{bmatrix} \in \mathbb{R}^{n\times 1}$

row vector $\begin{bmatrix}1 & 0.5 & 7 & 2\end{bmatrix} \in \mathbb{R}^{1\times n}$

Multiplying a matrix by a vector

K. N. Toosi
University of Technology

$$\begin{bmatrix} 1 & 2 & 7 & 4 \\ -1 & 0 & 2 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 10 \\ 20 \\ -10 \\ 30 \end{bmatrix} = \begin{bmatrix} 100 \\ 0 \\ 60 \end{bmatrix} = \begin{bmatrix} 1 \times 10 + 2 \times 20 + 7 \times (-10) + 4 \times 30 \\ (-1) \times 10 + 0 \times 20 + 2 \times (-10) + 1 \times 30 \\ 3 \times 10 + 2 \times 20 + 1 \times (-10) + 0 \times 30 \end{bmatrix}$$

$A \in \mathbb{R}^{3 \times 4}$ $u \in \mathbb{R}^4$ $v \in \mathbb{R}^3$

$$\begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} \times (10) + \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} \times (20) + \begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix} \times (-10) + \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix} \times (30)$$

linear combination of columns of A

Multiplying a matrix by a vector

K. N. Toosi
University of Technology

linear combination of columns of A

$$\underbrace{\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}}_{m \times 3} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} = \gamma_1 a_1 + \gamma_2 a_2 + \gamma_3 a_3 \in \mathbb{R}^m$$

$\mathbb{R}^{m \times 1}$

$a_i \in \mathbb{R}^m$

Define span in matrix form

K. N. Toosi
University of Technology

$$\operatorname{span}(u_1, u_2, \dots, u_n) = \{ \underbrace{\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n}_{\sum_{i=1}^{n} \alpha_i u_i} \mid \alpha_i \in \mathbb{R} \}$$

$$u_i \in \mathbb{R}^m$$

$$= \left\{ \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \;\middle|\; \alpha_i \in \mathbb{R} \right\} = \{ U a \mid a \in \mathbb{R}^n \}$$

$$U \in \mathbb{R}^{m \times n} \qquad a \in \mathbb{R}^n$$

where $U = [u_1, u_2, \dots, u_n]$

$$U = [u_1 \; u_2 \; \cdots \; u_n] \in \mathbb{R}^{m \times n}$$

Column space

$$\underline{\underline{U}} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \in \mathbb{R}^{m \times n}$$

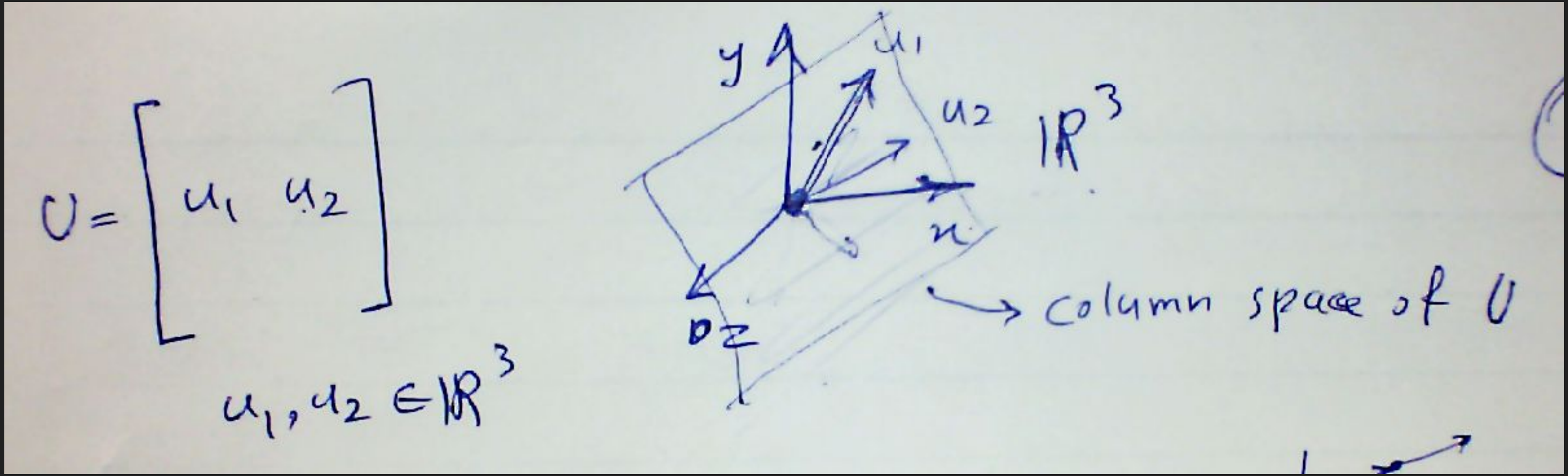
$$\text{span}(u_1, u_2, \ldots, u_n)$$

$= \text{column space of } U$

$= \text{range of } U$

$$= \{ Ua \mid a \in \mathbb{R}^n \}$$

Column space



Linear Subspace

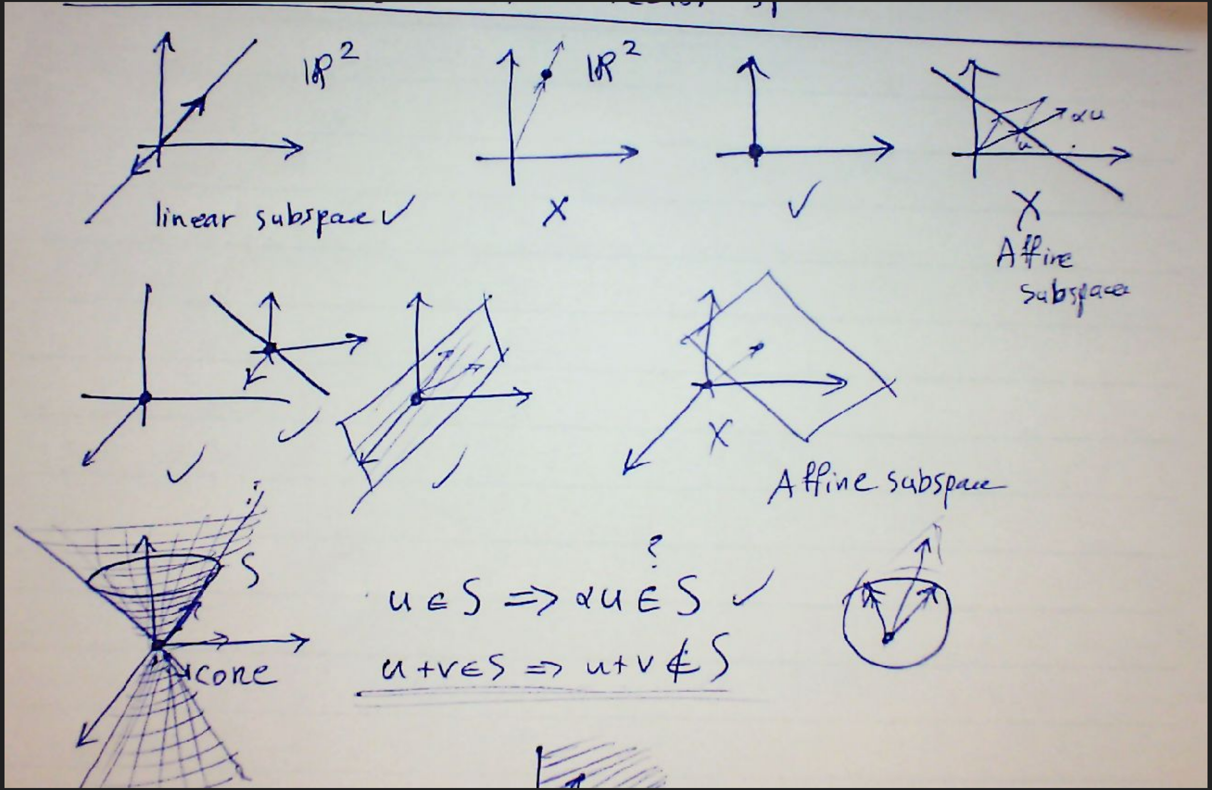
Let V be a vector space. A subset $S \subseteq V$ is called a linear subspace of V if

- For all $\mathbf{u} \in S$ and all $\alpha \in R \Rightarrow \alpha \mathbf{u} \in S$.
- For all $\mathbf{u}, \mathbf{v} \in S \Rightarrow \mathbf{u} + \mathbf{v} \in S$.

Alternatively,

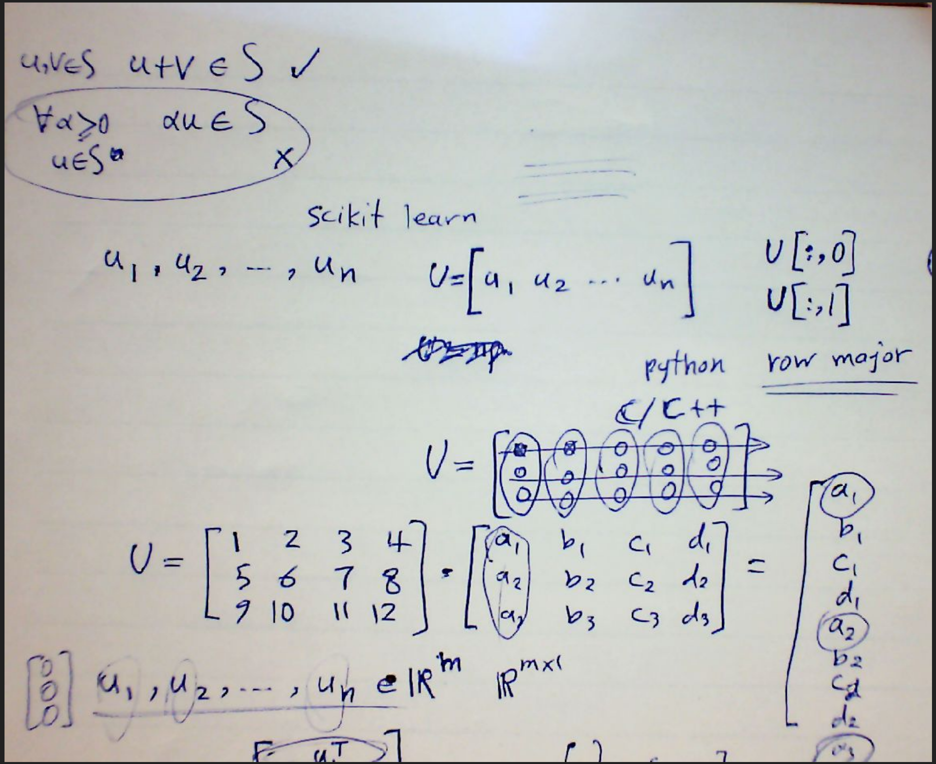
- For all $\mathbf{u}, \mathbf{v} \in S$ and $\alpha, \beta \in R \Rightarrow \alpha \mathbf{u} + \beta \mathbf{v} \in S$.

Linear Subspace

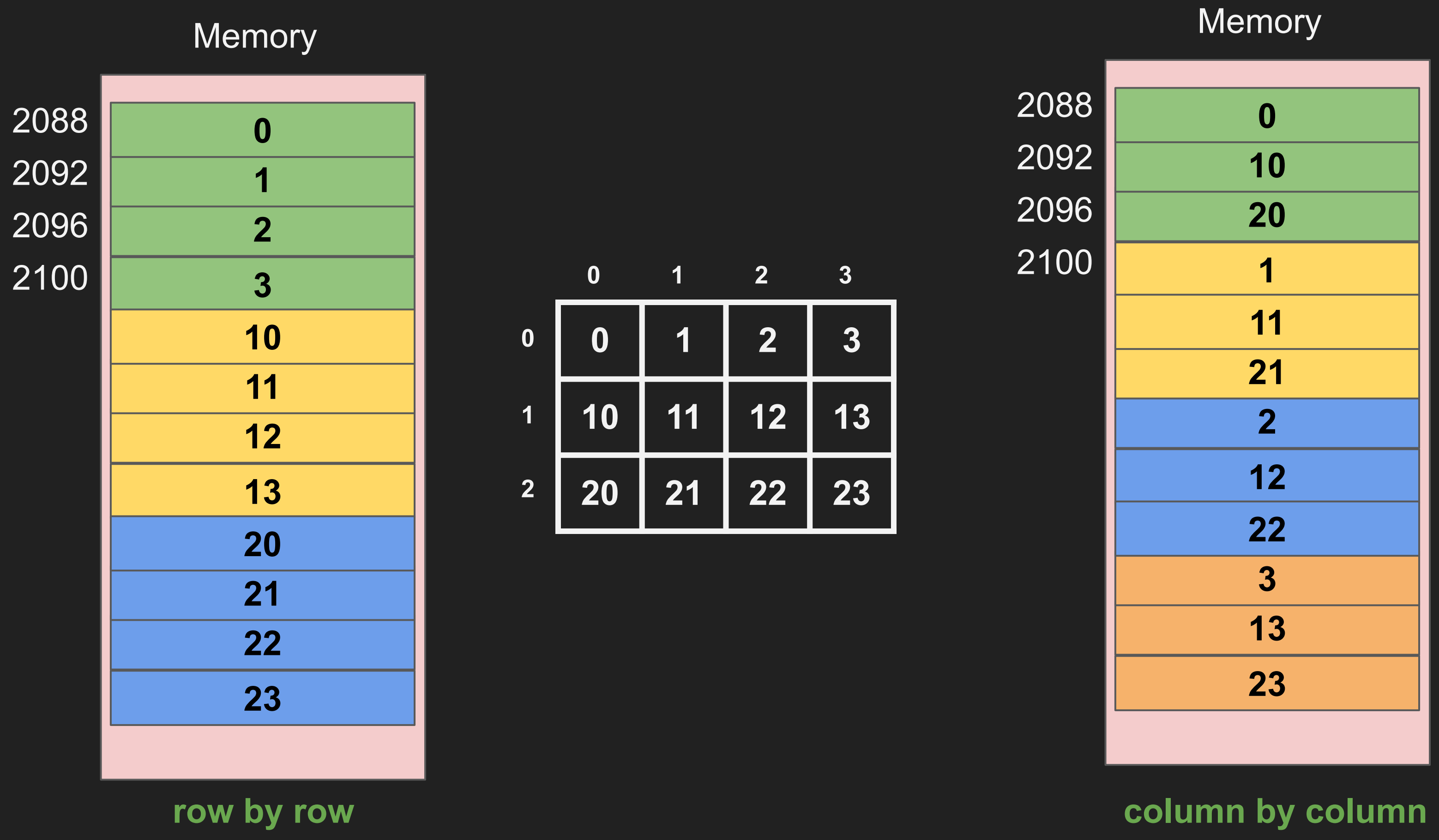


Data as columns of the matrix row-major programming languages

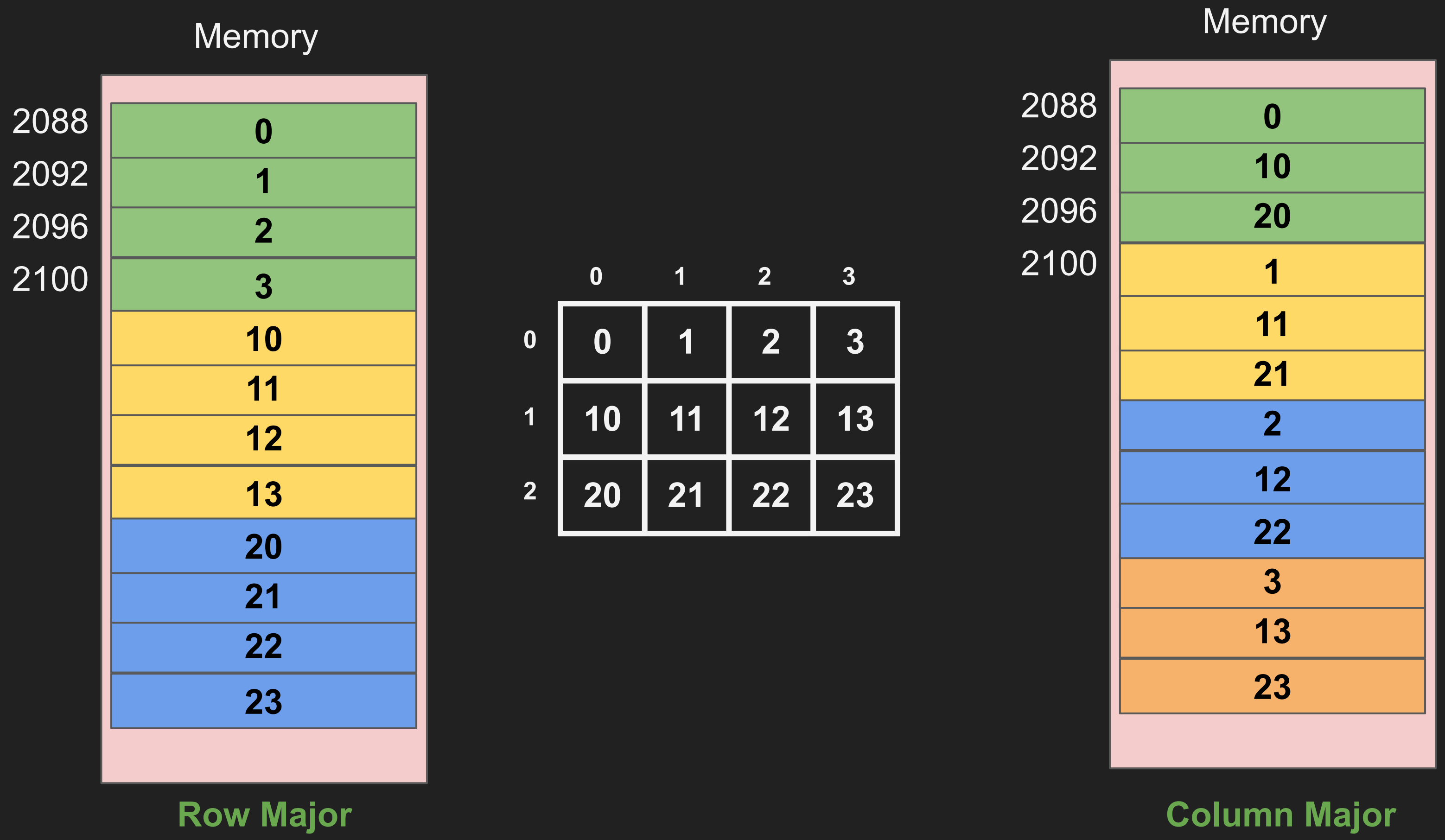
K. N. Toosi
University of Technology



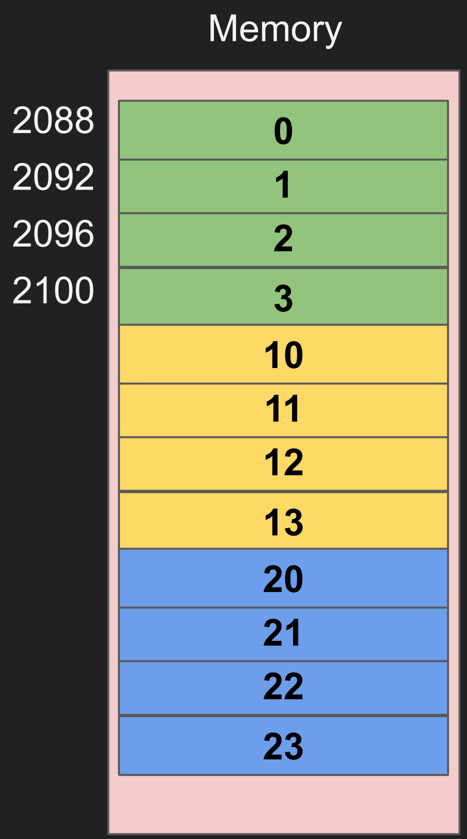
Representing Data and Data Batches



K. N. Toosi
University of Technology

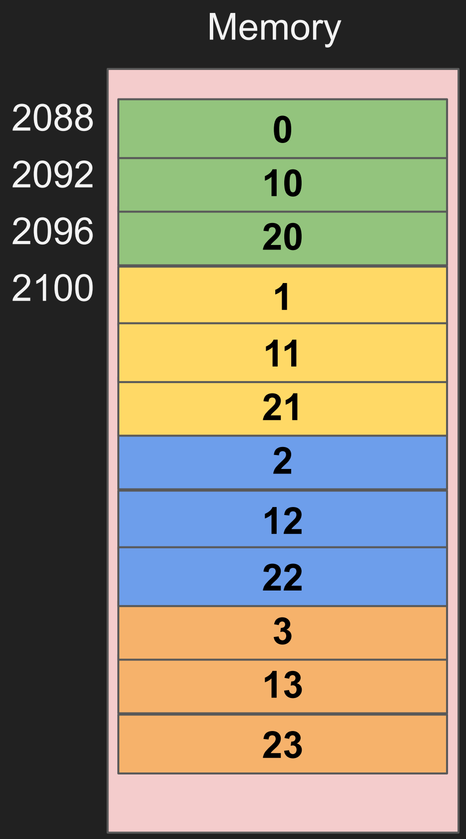


K. N. Toosi
University of Technology



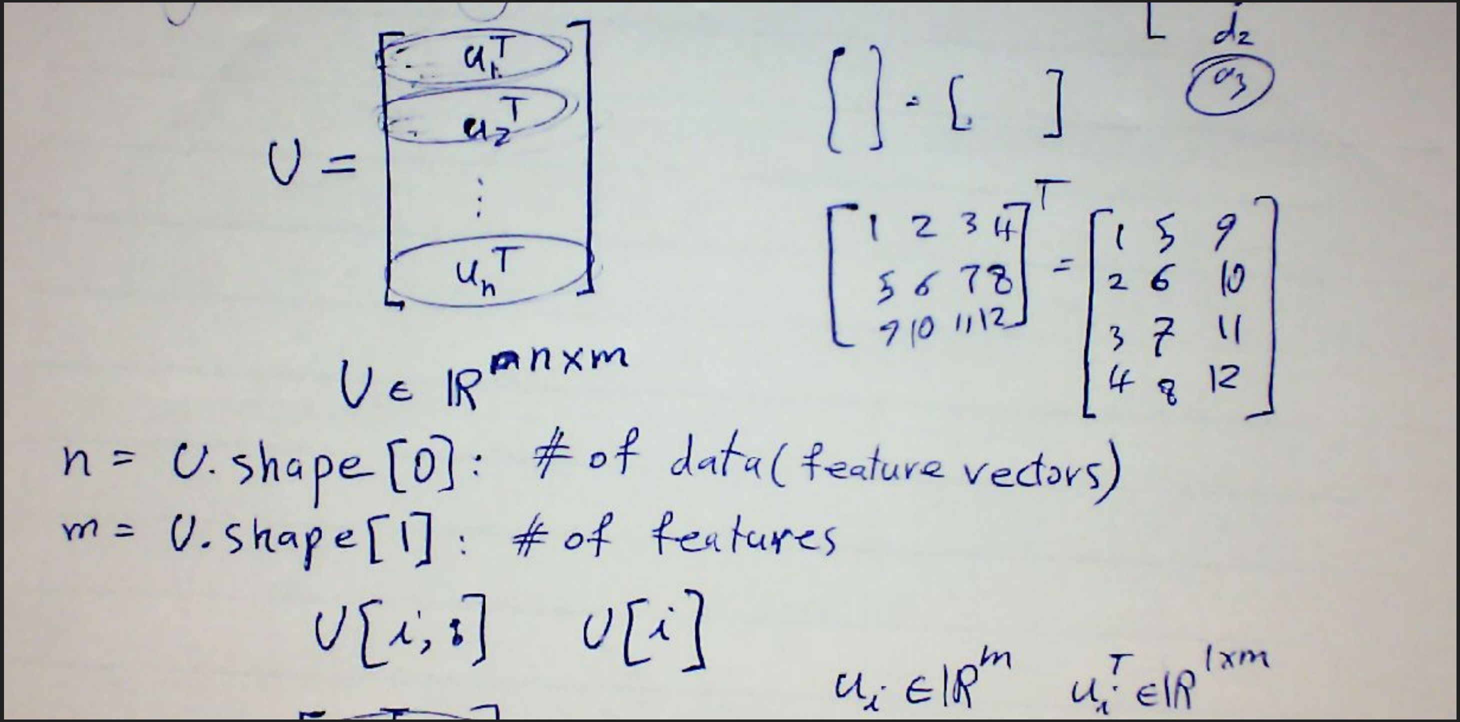
Row Major
(C,C++,Pascal, Python-numpy)

	0	1	2	3
0	0	1	2	3
1	10	11	12	13
2	20	21	22	23



Column Major
(Fortran, Matlab, R, …)

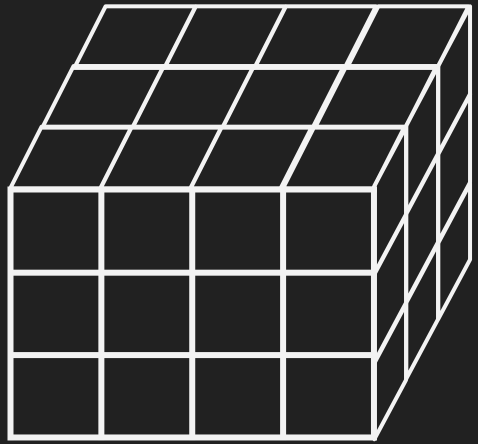
Data as rows of the matrix
row-major programming languages



ND Arrays: Row-major vs. Column-major

- What does row-major and column-major mean?
  - Matlab vs Numpy ND-arrays

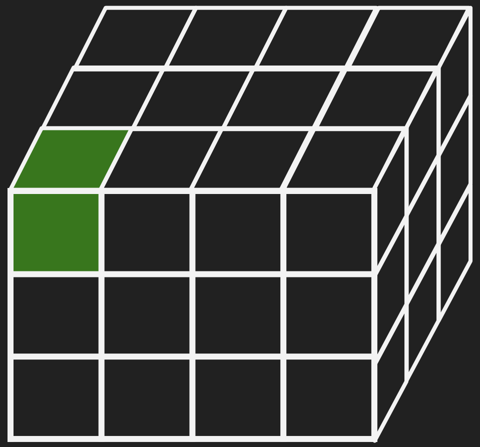
- **Row-major: last index moves fastest**



ND Arrays: Row-major vs. Column-major

- What does row-major and column-major mean?
  - Matlab vs Numpy ND-arrays

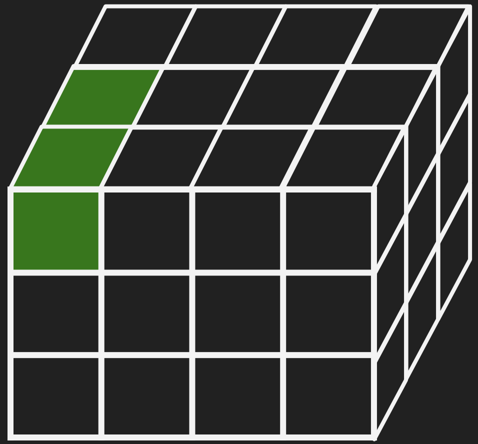
- **Row-major: last index moves fastest**



ND Arrays: Row-major vs. Column-major

- What does row-major and column-major mean?
  - Matlab vs Numpy ND-arrays

- **Row-major: last index moves fastest**

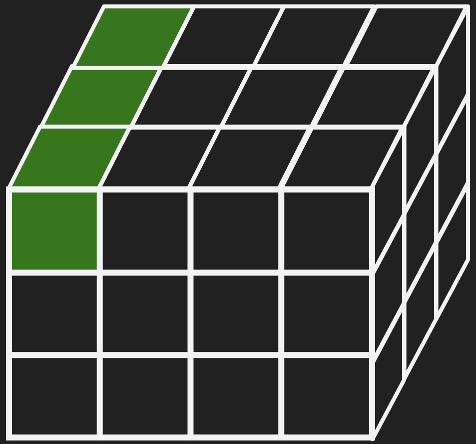


ND Arrays: Row-major vs. Column-major

K. N. Toosi
University of Technology

- What does row-major and column-major mean?
  - Matlab vs Numpy ND-arrays

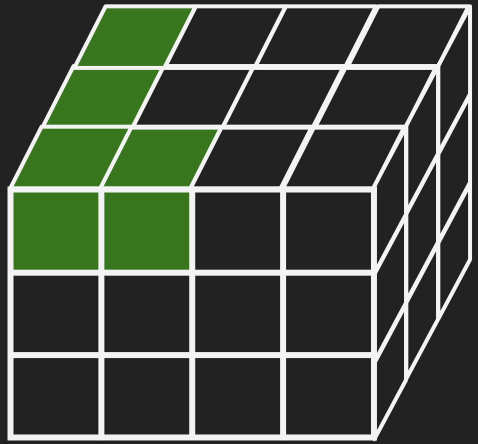
- Row-major: last index moves fastest



ND Arrays: Row-major vs. Column-major

K. N. Toosi
University of Technology

- What does row-major and column-major mean?
  - Matlab vs Numpy ND-arrays

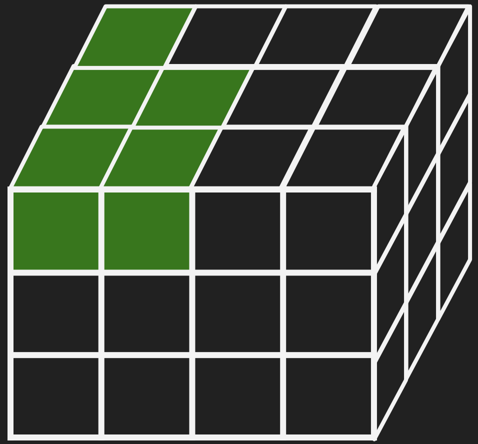


- **Row-major: last index moves fastest**

ND Arrays: Row-major vs. Column-major

- What does row-major and column-major mean?
  - Matlab vs Numpy ND-arrays

- **Row-major: last index moves fastest**

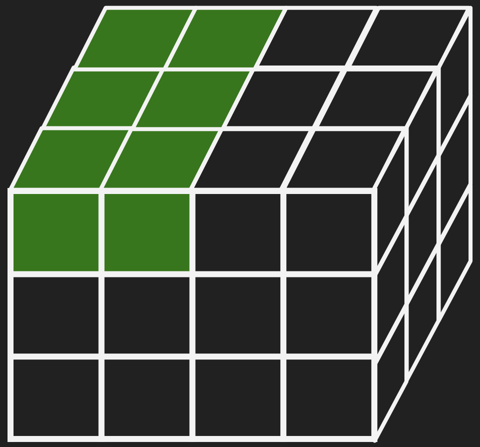


ND Arrays: Row-major vs. Column-major

K. N. Toosi
University of Technology

- What does row-major and column-major mean?
  - Matlab vs Numpy ND-arrays

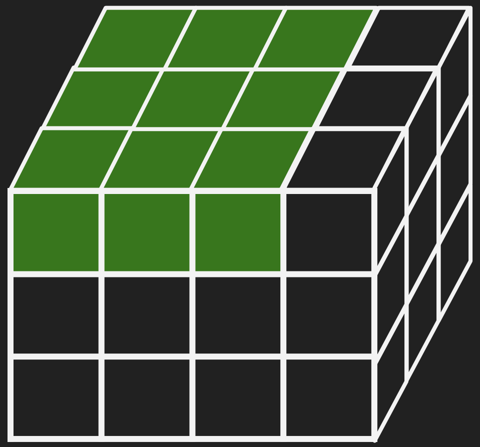
- **Row-major: last index moves fastest**



ND Arrays: Row-major vs. Column-major

- What does row-major and column-major mean?
  - Matlab vs Numpy ND-arrays

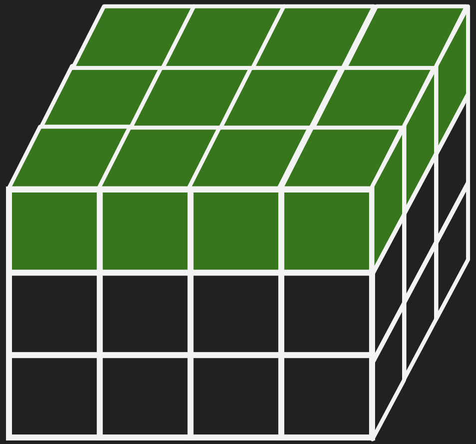
- **Row-major: last index moves fastest**



ND Arrays: Row-major vs. Column-major

- What does row-major and column-major mean?
  - Matlab vs Numpy ND-arrays

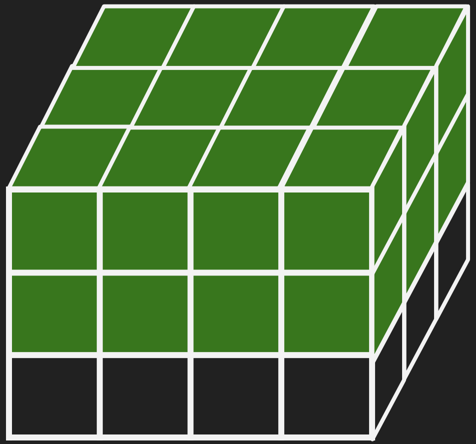
- **Row-major: last index moves fastest**



ND Arrays: Row-major vs. Column-major

- What does row-major and column-major mean?
  - Matlab vs Numpy ND-arrays

- **Row-major: last index moves fastest**

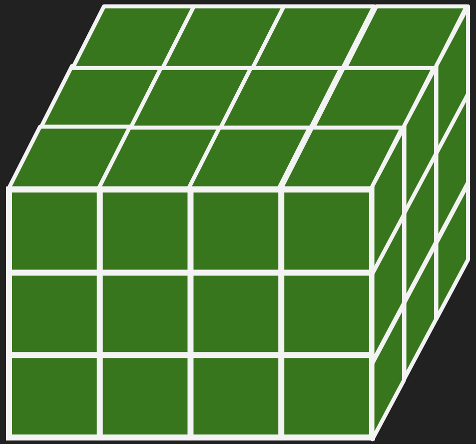


ND Arrays: Row-major vs. Column-major

K. N. Toosi
University of Technology

- What does row-major and column-major mean?
  - Matlab vs Numpy ND-arrays

- **Row-major: last index moves fastest**



Linear combination of matrix rows

K. N. Toosi
University of Technology

$U[i,:] \quad U[i]$

$u_i \in \mathbb{R}^{m} \quad u_i^T \in \mathbb{R}^{1\times m}$

$$\underbrace{\begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix}}_{\mathbb{R}^{1\times n}} \underbrace{\begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix}}_{\mathbb{R}^{n\times m}} = \underbrace{\alpha_1 u_1^T + \alpha_2 u_2^T + \cdots + \alpha_n u_n^T}_{\in \mathbb{R}^{1\times m}} \in \mathbb{R}^{1\times m}$$

Span of the rows of the matrix

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix} \begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \\ w_1 & w_2 \end{bmatrix} = \begin{bmatrix} \alpha_1 u_1 + \alpha_2 v_1 + \alpha_3 w_1 & \alpha_1 u_2 + \alpha_2 v_2 + \alpha_3 w_2 \end{bmatrix}$$

$$\alpha_1 \begin{bmatrix} u_1 & u_2 \end{bmatrix} + \alpha_2 \begin{bmatrix} v_1 & v_2 \end{bmatrix} + \alpha_3 \begin{bmatrix} w_1 & w_2 \end{bmatrix}$$

$$\alpha_1 u^T + \alpha_2 v^T + \alpha_3 w^T$$

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad v \in \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad w \in \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

$$\text{span}(u_1, u_2, \dots, u_n) = \left\{ \sum \alpha_i u_i \mid \alpha_i \in \mathbb{R} \;\; i = 1 \dots n \right\}$$

$$U = \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix} \qquad = \left\{ a^T U \mid a \in \mathbb{R}^n \right\}$$

$$a = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbb{R}^n \quad a^T = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix}$$

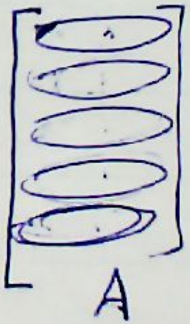
Row space of a matrix

The row space of a matrix $A \in \mathbb{R}^{m \times n}$

$= $ span of rows of A

$= \{ c^T A \mid c \in \mathbb{R}^m \}$ $[c_1 \; c_2 \; c_3 \; c_4 \; c_5]$

$\overset{?}{=} \{ A^T c \mid c \in \mathbb{R}^m \}$



Functions

K. N. Toosi
University of Technology

- Also maps, Transformations,
- What is a function?

functions / maps / tranformations

$f: X \to Y$

$\text{Domain}(f) = X$

$\text{Codomain}(f) = Y$

$\text{Range}(f) = \{ f(x) \mid x \in X \}$

Functions

functions / maps / tranformations

$f: X \longrightarrow Y$

$\text{Domain}(f) = X$

$\text{Codomain}(f) = Y$

$\text{Range}(f) = \{ f(x) \mid x \in X \}$

f: one-to-one (injective) $\quad f(x) = f(y) \Longrightarrow x = y$

یک به یک

f: onto (surjective) $\quad \text{Range}(f) = Y$

پوشا

$\forall y \in Y \;\; \exists x \in X : f(x) = y$

f: one-to-one & onto (bijective) $\Longrightarrow$ Invertible

یک به یک و پوشا $\quad$ معکوس پذیر

Functions

K. N. Toosi
University of Technology

f: one-to-one & onto (bijective) $\Longrightarrow$ Inrertible

یک به یک و پوشا — معکوس پذیر

bijective $\exists g$ such that ~~fog~~ $g(f(x)) = x \quad \forall x \in X$

$\exists g \; Y \to X$ $\qquad f(g(y)) = y \quad \forall y \in Y$

$g = f^{-1}$

Functions in linear algebra

- Here, we are interested in functions from a vector space V to a vector space U

$$(f: U \to V)$$

Linear Transformations

A linear map $f: V \longrightarrow U$

1. $f(u+v) = f(u)+f(v) \quad \forall u,v \in V$
2. $f(\alpha u) = \alpha f(u) \quad \forall u \in V, \alpha \in \mathbb{R}$, $\mathbb{C}$

$1,2 \iff f(\alpha u + \beta v) = \alpha f(u) + \beta f(v) \quad \forall u,v \in V \quad \forall \alpha,\beta \in \mathbb{R}$

$f(\alpha u)+f(\alpha v)$

A linear map preserves linear combinations

Linear Transformations

$f(u+v) = f(u) + f(v)$

$f(a\, u) = a\, f(u)$

does not matter if linear combination applied before or after transformation.