Computational Data Mining

Part 2: Linear Algebra
Vectors and Matrices

Instructor: Zahra Narimani

Department of Computer Science and Information Technology,
Institute for Advanced Studies in Basic Sciences, Zanjan, Iran

Inner product of vectors

$$(x, y) = x^T y$$

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 1 \cdot 1 + 2 \cdot 1 + 3 \cdot 1 + 4 \cdot 1 + 5 \cdot 1$$

Matrix-Vector Multiplication

A: m × n matrix

$$y = Ax, \qquad y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, \ldots, m.$$

```
for i=1:m
  y(i)=0;
  for j=1:n
    y(i)=y(i)+A(i,j)*x(j);
  end
end
```

$$\begin{pmatrix} \times \\ \times \\ \times \\ \times \end{pmatrix} = \begin{pmatrix} \leftarrow & - & - & \rightarrow \\ \leftarrow & - & - & \rightarrow \\ \leftarrow & - & - & \rightarrow \\ \leftarrow & - & - & \rightarrow \end{pmatrix} \begin{pmatrix} \uparrow \\ | \\ | \\ \downarrow \end{pmatrix}$$

Matrix-Vector Multiplication

A: m × n matrix

$$y = Ax, \qquad y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, \ldots, m.$$

```
for i=1:m
      y(i)=0;
end
for j=1:n
      for i=1:m
            y(i)=y(i)+A(i,j)*x(j);
      end
end
```

$$\begin{pmatrix} \uparrow \\ | \\ | \\ \downarrow \end{pmatrix} = \begin{pmatrix} \uparrow & \uparrow & \uparrow & \uparrow \\ | & | & | & | \\ | & | & | & | \\ \downarrow & \downarrow & \downarrow & \downarrow \end{pmatrix} \begin{pmatrix} \times \\ \times \\ \times \\ \times \end{pmatrix}$$

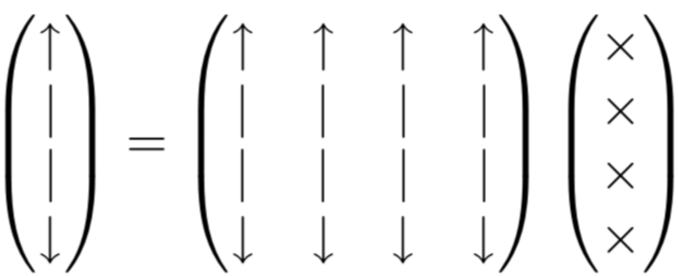
Matrix-Vector Multiplication

A: m × n matrix

$$y = Ax, \qquad y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, \ldots, m.$$

```
for i=1:m
    y(i)=0;
end
for j=1:n
    for i=1:m
        y(i)=y(i)+A(i,j)*x(j);
    end
end
```

```
y(1:m)=0;
for j=1:n
    y(1:m)=y(1:m)+A(1:m,j)*x(j);
end
```



Matrix-Matrix Multiplication

$$\mathbb{R}^{m \times n} \ni C = AB = (c_{ij}),$$

$$c_{ij} = \sum_{s=1}^{k} a_{is} b_{sj}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n.$$

Matrix-Matrix Multiplication

```
for i=1:m
  for j=1:n
    for s=1:k
      C(i,j)=C(i,j)+A(i,s)*B(s,j)
    end
  end
end
```

Matrix-Matrix Multiplication

```
for i=1:m
  for j=1:n
    for s=1:k
      C(i,j)=C(i,j)+A(i,s)*B(s,j)
    end
  end
end
```

Inner product version of the code:

```
for i=1:m
  for j=1:n
    C(i,j)=A(i,1:k)*B(1:k,j)
  end
end
```

Matrix-Matrix Multiplication – general form

```
for ...
  for ...
    for ...
      C(i,j)=C(i,j)+A(i,s)*B(s,j)
    end
  end
end
```

Matrix-Matrix product: column-oriented (SAXPY)

```
for j=1:n
  for s=1:k
    C(1:m,j)=C(1:m,j)+A(1:m,s)*B(s,j)
  end
end
```

Inner product of vectors

$$(x, y) = x^T y$$

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 1 \cdot 1 + 2 \cdot 1 + 3 \cdot 1 + 4 \cdot 1 + 5 \cdot 1$$

Vector Norms

- The size of a vector

$$\| x \|_1 = \sum_{i=1}^{n} |x_i|, \quad \text{1-norm,}$$

$$\| x \|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}, \quad \text{Euclidean norm (2-norm),}$$

$$\| x \|_\infty = \max_{1 \le i \le n} |x_i|, \quad \text{max-norm.}$$

- General p-norm

$$\| x \|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

Properties of vector norm

- Vector norm is a mapping from $R^n \rightarrow R$

$$\begin{aligned} \| x \| &\geq 0 \text{ for all } x, \\ \| x \| &= 0 \text{ if and only if } x = 0, \\ \| \alpha x \| &= |\alpha| \, \| x \|, \; \alpha \in \mathbb{R}, \\ \| x + y \| &\leq \| x \| + \| y \|, \text{ the triangle inequality.} \end{aligned}$$

Use norm for error approximation

- Absolute error

$$\| \delta x \| = \| \bar{x} - x \|$$

- Relative error

$$\frac{\| \delta x \|}{\| x \|} = \frac{\| \bar{x} - x \|}{\| x \|}$$

- In a finite dimensional vector space all vector norms are equivalent in the sense that for any two norms $\| \cdot \|_\alpha$ and $\| \cdot \|_\beta$ there exist constants m and M such that:

$$m \| x \|_\alpha \leq \| x \|_\beta \leq M \| x \|_\alpha$$

- Example ($x \in R^n$):

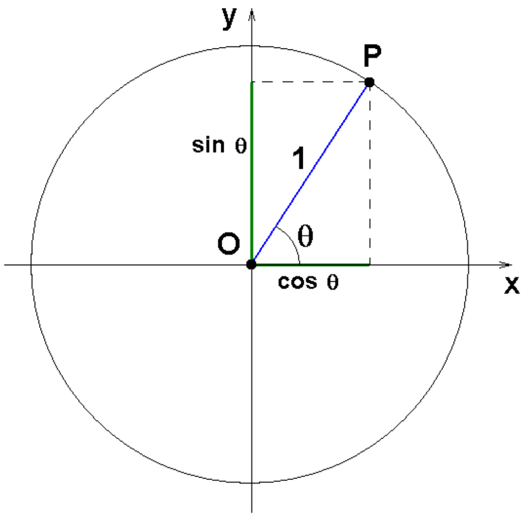
$$\| x \|_2 \leq \| x \|_1 \leq \sqrt{n} \| x \|_2$$

Cosine similarity

- Cosine of the angle between two vectors:

$$\cos\theta(x, y) = \frac{x^T y}{\| x \|_2 \; \| y \|_2}$$

- What if vectors are orthogonal?



Matrix Norms

- Let ||.|| be a vector norm, the corresponding matrix norm is defined as:

$$\| A \| = \sup_{x \neq 0} \frac{\| Ax \|}{\| x \|}$$

- Matrix norm satisfies (for $\alpha \in \mathbb{R}$):

$$\begin{aligned} \| A \| &\geq 0 \text{ for all } A, \\ \| A \| &= 0 \text{ if and only if } A = 0, \\ \| \alpha A \| &= |\alpha| \, \| A \|, \ \alpha \in \mathbb{R}, \\ \| A + B \| &\leq \| A \| + \| B \|, \text{ the triangle inequality.} \end{aligned}$$

Proposition 2.1. *Let* $\| \cdot \|$ *denote a vector norm and the corresponding matrix norm. Then*

$$\| Ax \| \leq \| A \| \, \| x \|,$$
$$\| AB \| \leq \| A \| \, \| B \|.$$

- Matrix 2-norm (matrix as an operator)
  - (square root of largest eigenvalue of A^TA)

$$\| A \|_2 = \left(\max_{1 \le i \le n} \lambda_i (A^T A) \right)^{1/2}$$

- Matrix 1-norm

$$\| A \|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|$$

- Matrix inf-norm

$$\| A \|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|$$

- Frobenius matrix norm
  - matrix as a point of space of dimension mn

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2}$$

$$\| A \|_F^2 = \mathrm{tr}(A^T A)$$

- Trace of matrix B = sum of diagonal elements

$$\mathrm{tr}(B) = \sum_{i=1}^{n} b_{ii}$$

Linear Independence: Bases

Given a set of vectors $(v_j)_{j=1}^n$ in $\mathbb{R}^m$, $m \geq n$,

$$\operatorname{span}(v_1, v_2, \ldots, v_n) = \left\{ y \mid y = \sum_{j=1}^{n} \alpha_j v_j \right\}$$

The vectors $(v_j)_{j=1}^n$ are called *linearly independent*

$\sum_{j=1}^n \alpha_j v_j = 0$ if and only if $\alpha_j = 0$ for $j = 1, 2, \ldots, n$.

Proposition 2.2. *Assume that the vectors* $(v_j)_{j=1}^n$ *are linearly dependent. Then some* v_k *can be written as linear combinations of the rest,*

$$v_k = \sum_{j \neq k} \beta_j v_j$$

Rank of Matrix

- Maximum number of linearly independent **column** vectors

- **Proposition 2.3.** An outer product matrix xy^T , where x and y are vectors in R^n, has rank 1.

$$xy^T = \begin{pmatrix} y_1 x & y_2 x & \cdots & y_n x \end{pmatrix} = \begin{pmatrix} x_1 y^T \\ x_2 y^T \\ \vdots \\ x_n y^T \end{pmatrix}$$

Nonsingular matrix

- A square matrix $A \epsilon \mathbb{R}^{n \times n}$ with rank n
- Has an inverse A^{-1} satisfying

$$A^{-1}A = AA^{-1} = \mathrm{I}$$

If we multiply **linearly independent vectors** by a **nonsingular matrix**, then the vectors remain linearly independent.

Proposition 2.4. *Assume that the vectors* $v_1, \ldots, v_p$ *are linearly independent. Then for any nonsingular matrix* T*, the vectors* $Tv_1, \ldots, Tv_p$ *are linearly independent.*

Proof. Obviously $\sum_{j=1}^{p} \alpha_j v_j = 0$ if and only if $\sum_{j=1}^{p} \alpha_j T v_j = 0$ (since we can multiply any of the equations by T or T^{-1}). Therefore the statement follows. $\square$

Weather prediction example

		today		
		sunny	cloudy	rainy
tomorrow	sunny	0.4	0.3	0.1
	cloudy	0.4	0.3	0.6
	rainy	0.2	0.4	0.3

If today is cloudy, what is the probability that tomorrow is:

- Cloudy?
- Sunny?
- Rainy?

Weather prediction example

		today		
		sunny	cloudy	rainy
tomorrow	sunny	0.4	0.3	0.1
	cloudy	0.4	0.3	0.6
	rainy	0.2	0.4	0.3

If today is cloudy, what is the probability that tomorrow is:

- Cloudy? 0.3
- Sunny? 0.3
- Rainy? 0.4

Weather prediction example

		today		
		sunny	cloudy	rainy
tomorrow	sunny	0.4	0.3	0.1
	cloudy	0.4	0.3	0.6
	rainy	0.2	0.4	0.3

If today is cloudy, what is the probability that the day after tomorrow is:

- Cloudy?
- Sunny?
- Rainy?

Any Question?