Lecture 3 – Linear Algebra Background

Dan Sheldon

September 17, 2012

Motivation

Preview of next class:

$$
\begin{aligned}
y^{(1)} &\approx w_0 + w_1 x_1^{(1)} + w_2 x_2^{(1)} + \ldots + w_d x_d^{(1)} \\
y^{(2)} &\approx w_0 + w_1 x_1^{(2)} + w_2 x_2^{(2)} + \ldots + w_d x_d^{(2)} \\
\ldots \\
y^{(N)} &\approx w_0 + w_1 x_1^{(N)} + w_2 x_2^{(N)} + \ldots + w_d x_d^{(N)}
\end{aligned}
$$

After linear algebra

$$\mathbf{y} \approx X\mathbf{w}$$

Linear Algebra in ML

Linear Algebra

- Succinct notation for models and algorithms
- Numerical tools (save coding!)

$$\mathbf{w} = (X^T X)^{-1} X^T \mathbf{y}$$

- Inspiration for new models and problems: Netflix

Netflix Movie Recommendations

	Gladiator	Silence of the Lambs	WALL-E	Toy Story
Alice	5	4	1	
Bob		5		2
Carol				5
David			5	5
Eve	5	4		

Matrix completion problem, matrix factorization

Today's Topics

- Matrices
- Vectors
- Matrix-Matrix multiplication (and special cases)
- Tranpose
- Inverse

Matrices

- A matrix is an rectangular array of numbers

$$A = \begin{bmatrix} 101 & 10 \\ 54 & 13 \\ 10 & 47 \end{bmatrix}$$

- When A has m rows and n columns, we say that:
  - A is an $m \times n$ matrix
  - $A \in \mathbb{R}^{m \times n}$
- The entry in row i and column j is denoted A_{ij}

Matrices

Example

$$A = \begin{bmatrix} 101 & 10 \\ 54 & 13 \\ 10 & 47 \end{bmatrix}$$

- $A \in \mathbb{R}^{3\times 2}$
- $A_{11} = 101$
- $A_{32} =$
- $A_{22} =$
- $A_{23} =$

Vectors

- A vector is an $n \times 1$ matrix:

$$\mathbf{x} = \begin{bmatrix} 8 \\ 2.4 \\ 1 \\ -10 \end{bmatrix}$$

- We write $\mathbf{x} \in \mathbb{R}^n$ (instead of $\mathbf{x} \in \mathbb{R}^{n\times 1}$)
- The ith entry is x_i

Vectors

Example

$$\mathbf{x} = \begin{bmatrix} 8 \\ 2.4 \\ 1 \\ -10 \end{bmatrix}$$

- $\mathbf{x} \in \mathbb{R}^4$
- $x_1 =$
- $x_4 =$

Addition

- If two matrices have the same size, we can add them by adding corresponding elements

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 3 & 5 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 7 \\ 2 & 4 \end{bmatrix}$$

- Subtraction is similar
- Matrices of different sizes *cannot be added or subtracted*

Scalar Multiplication

- A *scalar* $x \in \mathbb{R}$ is a real number (i.e., not a vector)

  e.g., 2, 3, π, $\sqrt{2}$, 1.843, $\ldots$

- Scalar times matrix:

$$2 \cdot \begin{bmatrix} 1 & 3 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ -4 & 0 \end{bmatrix}$$

  (multiply each entry by the scalar)

Matrix-Matrix Multiplication

- Can multiply two matrices *if their inner dimensions match*

$$A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}$$

$$C = AB \quad \in \mathbb{R}^{m \times p}$$

- The product has entries

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

Matrix-Matrix Multiplication

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

Move along ith row of A and jth row of B. Multiply corresponding entries, then add.

$$\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & {\color{red}c_{32}} & c_{33} \\ c_{41} & c_{42} & c_{43} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ {\color{red}a_{31}} & {\color{red}a_{32}} \\ a_{41} & a_{42} \end{bmatrix} \begin{bmatrix} b_{11} & {\color{red}b_{12}} & b_{13} \\ b_{21} & {\color{red}b_{22}} & b_{23} \end{bmatrix}$$

$${\color{red}c_{32} = a_{31} b_{12} + a_{32} b_{22}}$$

Matrix-Matrix Multiplication

Example

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}$$

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

Multiplication Properties

- Associative

$$(AB)C = A(BC)$$

- Distributive

$$A(B + C) = AB + AC$$

$$(B + C)D = BD + CD$$

- Not commutative

$$AB \neq BA$$

Matrix-Vector Multiplication

A (worthy) special case of matrix-matrix multiplication:

$$A \in \mathbb{R}^{m \times n}, \quad \mathbf{x} \in \mathbb{R}^n$$

$$\mathbf{y} = A\mathbf{x} \in \mathbb{R}^m$$

Definition

$$y_i = \sum_{j=1}^{n} A_{ij} x_j$$

Matrix-Vector Multiplication

$$y_i = \sum_{j=1}^{n} A_{ij} x_j$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$y_3 = a_{31} x_1 + a_{32} x_2$$

Matrix-Vector Multiplication

Example

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \mathbf{z} = \begin{bmatrix} 8 \\ 1.5 \end{bmatrix}$$

- $A\mathbf{x} = \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$
- $A\mathbf{z} = \begin{bmatrix} 6.5 \\ 4.5 \end{bmatrix}$

Transpose

Transposition of a matrix swaps the rows and columns

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix}, \quad A^T = \begin{bmatrix} 1 & 0 \\ -1 & 3 \end{bmatrix}.$$

Definition:

- Let $A \in \mathbb{R}^{m \times n}$
- The *transpose* $A^T \in \mathbb{R}^{n \times m}$ has entries

$$(A^T)_{ij} = A_{ji}.$$

Transpose

Example

$$A = \begin{bmatrix} 3 & 2 \\ -1 & 0 \\ 1 & 4 \end{bmatrix} \qquad A^T = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 0 & 4 \end{bmatrix}$$

Example

$$\mathbf{x} = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} \qquad \mathbf{x}^T = \begin{bmatrix} 1 & -3 & 2 \end{bmatrix}$$

Dot product

- A *special* special-case of matrix-matrix multiplication
- Let $\mathbf{x}, \mathbf{y}$ be vectors of same size ($\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$).
- Their dot product is

$$
\begin{aligned}
\mathbf{x}^T\mathbf{y} &= \sum_{i=1}^{n} x_i y_i \\
&= \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}
\end{aligned}
$$

Vector Norm

- The *norm* of a vector

$$\begin{aligned}\|\mathbf{x}\| &= \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} \\ &= \sqrt{\mathbf{x}^T\mathbf{x}}\end{aligned}$$

- Geometric interpretation: length of the vector

Transpose Properties

- Transpose of transpose

$$(A^T)^T = A$$

- Transpose of sum

$$(A + B)^T = A^T + B^T$$

- Transpose of product

$$(AB)^T = B^T A^T$$

Identity

- The identity matrix $I \in \mathbb{R}^{n\times n}$ has entries

$$I_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases},$$

$$I_{1\times 1} = [1], \qquad I_{2\times 2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad I_{3\times 3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

- For any A, B of appropriate dimensions

$$IA = A$$
$$BI = B$$

Inverse

- The inverse $A^{-1} \in \mathbb{R}^{n \times n}$ of a *square* matrix $A \in \mathbb{R}^{n \times n}$ satisfies
$$AA^{-1} = I = A^{-1}A$$
- Compare to division of scalars
$$xx^{-1} = 1 = x^{-1}x$$
- Not all matrices are invertible
  - E.g., A not square, $A = [0]$, $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, many more

Inverse

Example

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

Is B the inverse of A?

Example

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix} \qquad A^{-1} = \begin{bmatrix} 1 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{bmatrix}$$

Verify on your own.

Inverse Properties

- Inverse of inverse
$$(A^{-1})^{-1} = A$$
- Inverse of product
$$(AB)^{-1} = B^{-1}A^{-1}$$
- Inverse of transpose
$$(A^{-1})^T = (A^T)^{-1} := A^{-T}$$

What You Should Know

- Definitions of matrices and vectors
- Meaning of matrix multiplication
  - Systems of equations $\longrightarrow$ matrix-vector equations
- Properties of multiplication
- Properties of inverse, transpose
  - Get familiar with these as course goes on