The Identity matrix

Remark: we'll see this matrix again. It's called the identity. But for now its just all the right hand sides bunched together...

$$I = \begin{bmatrix} e_1 \,|\, e_2 \,|\, \cdots \,|\, e_n \end{bmatrix} \in \mathbb{R}^{n \times n}$$

$e_1, e_2, \ldots, e_n \in \mathbb{R}^n$

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots\ldots, \quad e_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

Why is I called the identity... because that matrix corresponds to the identity function $f(x) = x$.

That is $f(x) = Ix$

Example: $n = 3$, $I \in \mathbb{R}^{3 \times 3}$

$$I = \begin{bmatrix} e_1 \,|\, e_2 \,|\, e_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Use matrix–vector multiplication to find Ix

$$Ix = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Row representation

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} [1\ 0\ 0] \cdot (x_1, x_2, x_3) \\ [0\ 1\ 0] \cdot (x_1, x_2, x_3) \\ [0\ 0\ 1] \cdot (x_1, x_2, x_3) \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Column representation

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

2.3 Characterizations of Invertible Matrices

The Invertible Matrix Theorem

Let A be a square $n \times n$ matrix. Then the following statements are equivalent. That is, for a given A, the statements are either all true or all false.

✓a. A is an invertible matrix.

✓b. A is row equivalent to the $n \times n$ identity matrix. ← This was actually the algorithm for finding the inverse..

✓c. A has n pivot positions. ← matrix is square so this means pivot in every column and pivot in every row. Thus $Ax = b$ has a unique solution x for every b.

✓d. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

✓e. The columns of A form a linearly independent set.

✓f. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.

g. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

h. The columns of A span $\mathbb{R}^n$.

i. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.

j. There is an $n \times n$ matrix C such that $CA = I$.

k. There is an $n \times n$ matrix D such that $AD = I$.

l. A^T is an invertible matrix.

every column and

(d) The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

This mean A has no free variables (otherwise lots of solutions) So there is a pivot in every column. Since A is square, there is a pivot in every row. So A is row equivalent to the $n \times n$ identity matrix.

e. The columns of A form a linearly independent set.

Means there are no free variables. Same as above.. That is if the columns were dependent then $Ax=0$ would have a non-trivial solution given by the dependency relation...

f. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.

Given x let $b = Ax$. Then one-to-one means that $Ax = b$ has only one solution.., so no free variables.

Note, the whole idea is that since the matrix is square that a pivot in every column implies there is a pivot in every row ... and vice versa...

This one seems difficult...

l. A^T is an invertible matrix.

Some observations about transposes... matrix-vector products and dot products...

$x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$ then $x \cdot y = \sum_{i=1}^{n} x_i y_i$

$x \in \mathbb{R}^3$ and $y \in \mathbb{R}^3$ then $x \cdot y = x_1 y_1 + x_2 y_2 + x_3 y_3$

when we write this we mean

$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ we also write this as $x = (x_1, x_2, x_3)$ to save space...

This is different than $[x_1 \;\; x_2 \;\; x_3] = x^T$

a row vector but can also think of this as a matrix

$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$

$[x_1 \;\; x_2 \;\; x_3][x \cdot y]$

Thus, $x^T y = \boxed{x \cdot y}$

What about A^T when $A \in \mathbb{R}^{m \times n}$

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad A^T = \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{n \times m}$$

What does A^T mean in linear algebra?

$y = Ax \qquad x \in \mathbb{R}^n \qquad y \in \mathbb{R}^m \qquad A \in \mathbb{R}^{m \times n}$

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

$$y^T z = y \cdot z = \sum_{i=1}^{m} y_i z_i = z \cdot y = z^T y$$

$$(Ax)^T z = z^T A x$$

$$A = \begin{bmatrix} a_1 \,\Big|\, a_2 \,\Big|\, \cdots \,\Big|\, a_n \end{bmatrix} \qquad A^T = \begin{bmatrix} a_1^T \\ \hline a_2^T \\ \hline \vdots \\ \hline a_n^T \end{bmatrix}$$

$$Ax = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n$$

$$(Ax)^T = (x_1 a_1 + x_2 a_2 + \cdots + x_n a_n)^T$$

$$= x_1 a_1^T + x_2 a_2^T + \cdots + x_n a_n^T$$

$$(Ax)^T z = (x_1 a_1^T + x_2 a_2^T + \cdots + x_n a_n^T)\, z$$

$$= x_1 a_1^T z + x_2 a_2^T z + \cdots + x_n a_n^T z$$

On the other hand

$$A^T z = \begin{bmatrix} a_1^T \\ \hline a_2^T \\ \hline \vdots \\ \hline a_n^T \end{bmatrix} z = \begin{bmatrix} a_1 \cdot z \\ a_2 \cdot z \\ \vdots \\ a_n \cdot z \end{bmatrix} = \begin{bmatrix} a_1^T z \\ a_2^T z \\ \vdots \\ a_n^T z \end{bmatrix}$$

Therefore

$$Ax \cdot z = (Ax)^T z = x \cdot \begin{bmatrix} a_1^T z \\ a_2^T z \\ \vdots \\ a_n^T z \end{bmatrix} = x \cdot A^T z$$

and so

$$Ax \cdot z = x \cdot A^T z$$

So the transpose is related to moving the matrix A over the dot product