LINEAR ALGEBRA INTRO

1. Linear algebra

1.1. Basics.

- To truly understand regression, we need to learn a bit of linear algebra
- We begin with some notation. A **vector** is a linear of numbers

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

- A $n \times m$ **matrix** is an array of numbers

$$A = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix}$$

  We can form an $n \times m$ matrix by taking a set of m n-vectors $\boldsymbol{a}_1, \dots \boldsymbol{a}_n$ and using them as columns in our matrix. We write this as

$$A = \begin{bmatrix} | & | & \cdots & | \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_2 \\ | & | & \cdots & | \end{bmatrix}$$

  We could do the same thing with rows. I'll note that we will mostly work with $n \times n$ matrices, so if you struggle with all the notation, you imagine $m = n$.
- The basic operations we can perform on matrices are
  - Addition: We add the elements element wise

$$\begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} + \begin{bmatrix} b_{11} & \dots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nm} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & \dots & a_{1m}+b_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & \dots & a_{nm}+b_{nm} \end{bmatrix}$$

    For example,

$$\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} + \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ 5 & 10 \end{bmatrix}$$

  - Multiplication by a scale: We multiply each element by the scale

$$c \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} = \begin{bmatrix} ca_{11} & \dots & ca_{1m} \\ \vdots & \ddots & \vdots \\ ca_{n1} & \dots & ca_{nm} \end{bmatrix}$$

    For example

$$5 \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 15 & 25 \end{bmatrix}$$

  - Transpose: We exchange the columns and rows:

$$\begin{bmatrix} | & | & \cdots & | \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_2 \\ | & | & \cdots & | \end{bmatrix}^T = \begin{bmatrix} — & \boldsymbol{a}_1^T & — \\ — & \boldsymbol{a}_2^T & — \\ & \vdots & \\ — & \boldsymbol{a}_n^T & — \end{bmatrix}$$

    Taking the transpose of an $n \times m$ matrix produces an $m \times n$ matrix. For example

$$\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}$$

    Note that a vector is matrix and

$$\begin{bmatrix} 1 \\ 3 \end{bmatrix}^T = [1, 3]$$

    The left hand side is what we call a row vector.
  - Multiplication. We first define multiplication of vectors. We can multiple the transpose of an n vector by an n vector (these are vectors)

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}^T \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = a_1 b_2 + a_1 b_2$$

We we define matrix multiplication of an $n \times m$ matrix by an $m \times k$ matrix. The product is obtained by multiplying the columns of the matrix on the left by the rows of the matrix on the right. If the vectors $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ and $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ are n-vectors, then

$$\begin{bmatrix} — & \boldsymbol{b}_1^T & — \\ — & \boldsymbol{b}_2^T & — \\ & \vdots & \\ — & \boldsymbol{b}_n^T & — \end{bmatrix} \begin{bmatrix} | & | & \cdots & | \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_k \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1^T \boldsymbol{a}_1 & \ldots & \boldsymbol{b}_1^T \boldsymbol{a}_k \\ \vdots & \ddots & \vdots \\ \boldsymbol{b}_n^T \boldsymbol{a}_1 & \ldots & \boldsymbol{b}_n^T \boldsymbol{a}_k \end{bmatrix}$$

where $\boldsymbol{b}_1$ are the rows of B. For example

$$\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 \times 1 + 2 \times 2 & 3 \times 1 + 5 \times 2 \\ 3 \times 1 + 2 \times 5 & 3 \times 3 + 5 \times 5 \end{bmatrix} = \begin{bmatrix} 5 & 13 \\ 13 & 34 \end{bmatrix}.$$

- We end by introducing the **standard basis vectors** $\boldsymbol{e}_i$ defined as the n-vector where all components are zero except the ith one. Usually n is whatever dimension we are working in, e.g. for a system with two equations $n = 2$. The matrix whose columns are the basis vectors is the Identity matrix:

$$I = \begin{bmatrix} | & | & \cdots & | \\ \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_n \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}$$

To test your understanding, you should try to prove that $I\boldsymbol{a} = \boldsymbol{a}$ for any vector n where I is the n-dimensional identity matrix.