These are brief notes for the lecture on Wednesday September 8, 2010: they are not complete, but they are a guide to what I want to say today. They are not guaranteed to be correct.

2.2. Matrix Operations, continued

Transpose of a matrix

Recall that the transpose of a $m \times n$ matrix A is the matrix A^T having (i,j)-entry a_{ji}.

For example, $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ has transpose $A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$. Thus the rows of A become the columns of A^T and vice versa.

Theorem 3. *Let A and B be matrices whose sizes are appropriate for the following sums and products to be defined*

1. $(A^T)^T = A$
2. $(A+B)^T = A^T + B^T$.
3. *For any scalar r, $(rA)^T = rA^T$.*
4. $(AB)^T = B^T A^T$

Note: The transpose of AB reverses the order of the product of the transposes: in general, if the product AB exists, then it means that the number of columns of A is equal to the number of rows of B: however, usually the product $A^T B^T$ will not exist! Even when it does exist, it is usually not equal to $(AB)^T$.

For example: if $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, and $B = \begin{pmatrix} 5 & 1 & -1 \\ 1 & 2 & 2 \end{pmatrix}$ then

$$AB = \begin{pmatrix} 7 & 5 & 3 \\ 9 & 11 & 5 \end{pmatrix} \qquad (AB)^T = \begin{pmatrix} 7 & 9 \\ 5 & 11 \\ 3 & 5 \end{pmatrix} = B^T A^T$$

but A^T is 2×2 and B^T is 3×2, so $A^T B^T$ isn't even defined.

Let's look at this further: if A and B are such that AB exists, and $A^T B^T$ exists, and A is $m \times n$ then B has n rows (since AB exists) and m columns (since $A^T B^T$ exists). Now, AB is an $m \times m$ matrix, so $(AB)^T$ is $m \times m$, whereas $A^T B^T$ is $n \times n$.

Even when A and B are both square matrices and have the same sizes (so it could be possible that $(AB)^T = A^T B^T$) it usually doesn't happen. Almost any random example you write down of two 2×2 matrices will have $(AB)^T \neq A^T B^T$.

Example

2.3. The Inverse of a Matrix

Suppose we have a linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$: we will say that the linear transformation has an inverse transformation if for every $\underline{b} \in \mathbb{R}^m$ there is exactly one $\underline{x} \in \mathbb{R}^n$ so that $T(\underline{x} = \underline{b}$. Note that if T is invertible, this means that T is onto (every equation can be solved: hence $m \leq n$) and T is 1-1 (every equation has at most one solution: hence $n \leq m$).

Hence an invertible linear transformation has a matrix which must be square. Which square matrices are invertible? What does it mean for a square matrix to be invertible?

Suppose that $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is an invertible linear transformation: then we can define $S : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ so that $T\underline{x} = \underline{u}$ if and only if $\underline{x} = S\underline{u}$. Furthermore, for every vector $\underline{x} \in \mathbb{R}^n$, $S(T(\underline{x})) = \underline{x}$, and for every $\underline{u} \in \mathbb{R}^n$, $T(S(\underline{u})) = \underline{u}$.

It turns out that S must also be linear: we'll assume that $T(\underline{x}) = \underline{u}$ and $T(\underline{y}) = \underline{v}$. Then $S(\underline{u}) = \underline{x}$ and $S(\underline{v}) = \underline{y}$.

We'll show that $S(ru) = rS(\underline{u})$. Indeed, $S(T(r\underline{x})) = r\underline{x}$, so we get

$$S(r\underline{u}) = S(rT(\underline{x})) = S(T(r\underline{x})) = r\underline{x} = rS(\underline{u})$$

so that S commutes with scalar addition.

Likewise,

$$S(\underline{u} + \underline{v}) = S(T(\underline{x}) + T(\underline{y})) = S(T(\underline{x} + \underline{y})) = (\underline{x} + \underline{y}) = S(\underline{u}) + S(\underline{v})$$

so that S commutes with addition.

So we see that if T is an invertible linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$, so is S. Hence we can represent T by a square matrix A and S by a square matrix B.

Then $S(T(\underline{x})) = \underline{x}$ for all $\underline{x}$ means that $BA\underline{x} = \underline{x}$ for every $\underline{x}$. In particular, if $C = BA$, then we have $C\underline{e}_j = \underline{e}_j$, so that we obtain that C must be the identity matrix I_n.

Similarly, $T(S(\underline{u})) = \underline{u}$ for every $\underline{u}$, and hence $AB = I_n$ is also the identity matrix.

Definition. *An $n \times n$ matrix is* invertible *if there exists a $n \times n$ matrix B so that $AB = I_n = BA$. B is called the* inverse *of A and is denoted by A^{-1}. A matrix which is not invertible is said to be* singular.

For 2×2 matrices we have a very simple form for the inverse.

DEFINITION. *Let* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. *We define (and this only works for* 2×2 *matrices) the determinant of* A *to be the quantity* $\det(A) = ad - bc$.

THEOREM 4. *Let* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. *Then* A *is invertible if and only if* $\det(A)$ *is non-zero, in which case*

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

If $\det(A) = 0$ *then* A *is* singular.

Proof:

THEOREM 5. *If* A *is an invertible* $m \times m$ *matrix, then for every* $\underline{b} \in \mathbb{R}^n$, *the equation* $A\underline{x} = \underline{b}$ *has a unique solution, namely* $\underline{x} = A^{-1}\underline{b}$.

Proof:

THEOREM 6.

(1) *If A is invertible, then so is A^{-1}, and its inverse is given by*
$$(A^{-1})^{-1} = A$$
.

(2) *If A and B are invertible $n \times n$ matrices (that is, they are both invertible and they are the same size) then AB is invertible, and*
$$(AB)^{-1} = B^{-1}A^{-1}.$$

(3) *If A is invertible, then so is A^T, and its inverse is given by*
$$(A^T)^{-1} = (A^{-1})^T.$$

Proof: