What is a Matrix?

An n by m matrix over $\mathbb{R}$ is a box of numbers, consisting of n rows and m columns.

$$\begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix}$$

where $x_{ij} \in \mathbb{R}$. For example:

- $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 3 \end{pmatrix}$ is a 2 by 3 matrix.
- $\begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 2 & 3 \end{pmatrix}$ is a 3 by 2 matrix.
- $\begin{pmatrix} 1 & 2 & 0 & 1 \end{pmatrix}$ is a 1 by 4 matrix (sometimes called a **row vector**)
- $\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ is a 3 by 1 matrix (sometimes called a **column vector**)
- $(\cdot)$ is 0 by 0 matrix (and a 5 by 0 matrix, and a 0 by 7 matrix...)

A matrix over $\mathbb{C}^n$ is then just a box with complex numbers.

What is a Matrix good for?

- They can be a natural way to present multiple similar lists.
- Wins, Losses, Ties for a whole bunch of teams.
- They might represent lists of probabilities, except for different people, or situations.
- They might describe a bunch of different points in space.
- They might describe the number of different direct daily flights between cities.
- They could describe the table of values for a function, even a function of multiple variables.
- They could describe the information contained when someone gives you a system of linear equations.
- They could describe a process for systematically making changes to other boxes of numbers! We often describe systematic ways of modifying column vectors this way!!!!

A box of numbers could be any box of numbers.
Just like with vectors, it is always a good idea when you have a box of numbers, to remember why you have a box of numbers.
If a box of numbers represents information about airline flights, it might not be so useful for planning sports bets.

What can I do with a box of numbers?

Just like with vectors, depending on what your box of numbers represents, most random things you can do with them are pretty meaningless.
So just like with vectors, it is useful to focus on doing things which have meaning for lots of different types of boxes of numbers.
But just as before, the context will effect the meaning you should associate to any of these operations.

So what are we going to do with boxes of numbers?

I can rescale the entire box of numbers by some constant.

$$C\begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix} = \begin{pmatrix} Cx_{11} & \cdots & Cx_{1m} \\ \vdots & \ddots & \vdots \\ Cx_{n1} & \cdots & Cx_{nm} \end{pmatrix}$$

which is just like we did with vectors.

The ij entry will be Cx_{ij}.

$$7\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 7 & 14 \\ 21 & 28 \\ 35 & 42 \end{pmatrix}$$

So what are we going to do with boxes of numbers?

I can add together boxes of numbers of the same size.

$$\begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix} + \begin{pmatrix} y_{11} & \cdots & y_{1m} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nm} \end{pmatrix} = \begin{pmatrix} x_{11} + y_{11} & \cdots & x_{1m} + y_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} + y_{n1} & \cdots & x_{nm} + y_{nm} \end{pmatrix}$$

which is just like we did with vectors.

- This only works if the sizes are exactly the same.
- The *ij* entry of the sum is $x_{ij} + y_{ij}$.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} + \begin{pmatrix} 7 & 6 \\ 4 & 5 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 8 & 8 \\ 7 & 8 \\ 8 & 7 \end{pmatrix}$$

So what are we going to do with boxes of numbers?

I can arbitrarily glue them together.
Imagine A is nxm and B is ℓxk, I can always do things like:

$$\begin{pmatrix} A & 0_{nxk} \\ 0_{\ell xm} & B \end{pmatrix}$$

which will be an $n+\ell$ by $m+k$ matrix.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 7 & 8 & 9 \\ 4 & 3 & 2 \end{pmatrix}$$

would become

$$\begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 \\ 5 & 6 & 0 & 0 & 0 \\ 0 & 0 & 7 & 8 & 9 \\ 0 & 0 & 4 & 3 & 2 \end{pmatrix}$$

This is in general an advanced thing to do, it IS the equivalent of concatenating vectors.
It is also sometimes a useful notational thing to do.

Advanced operations

- What if I want to make a new box of numbers that combines the first two rows (but keeps the rest the same).?

$$\begin{pmatrix}1&2&3\\4&3&2\\1&5&7\end{pmatrix} \Rightarrow \begin{pmatrix}5&5&5\\1&5&7\end{pmatrix} = \begin{pmatrix}1&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&2&3\\4&3&2\\1&5&7\end{pmatrix}$$

- What about the first two columns What if I want to make a new box of numbers that combines the first two rows (but keeps the rest the same).?

$$\begin{pmatrix}1&2&3\\4&3&2\\1&5&7\end{pmatrix} \Rightarrow \begin{pmatrix}3&3\\7&2\\6&7\end{pmatrix} = \begin{pmatrix}1&2&3\\4&3&2\\1&5&7\end{pmatrix}\begin{pmatrix}1&0\\1&0\\0&1\end{pmatrix}$$

- What if I want to delete the first row and first column?

$$\begin{pmatrix}1&2&3\\4&3&2\\1&5&7\end{pmatrix} \Rightarrow \begin{pmatrix}3&2\\5&7\end{pmatrix} = \begin{pmatrix}0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&2&3\\4&3&2\\1&5&7\end{pmatrix}\begin{pmatrix}0&0\\1&0\\0&1\end{pmatrix}$$

These are the sorts of things that matrix multiplication does.

Matrix multiplication in different contexts describes a huge number of interesting things you might want to do to a box of numbers.

There are many ways of interpreting what it is doing.

Interprettation of Matrix Multiplication

Matrix multiplication is a process by which if we have A is nxm and B is mxℓ, then we can create an nxℓ boxes

$$A \cdot B$$

We have just seen that:

- You can interpret this as the matrix A describing a bunch of modifications to make to B.
- You can interpret this as the matrix B describing a bunch of modifications to make to A.

But you could also iterpret this as:

- If A describes some changes I want to things, X, by the process AX, and B describes changes I might also want to make to things by BX then

  $$AB$$

  describes the changes I would get to thing X by first doing B, then doing A. So AB gives a single set of modification rules that corresponds to doing one and then the other.

Matrix multipliation is amazingly versatile for having different interpretations in different contexts, we will see some of these on Bonus questions on assignments!

Interprettation of Matrix Multiplication

If A is a matrix, what does:

$$\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} A$$

mean?

the result is the first row of A plus twice the second row, plus three times the third!

$$\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 4 & 5 \end{pmatrix} = 1 \begin{pmatrix} 1 & 2 \end{pmatrix} + 2 \begin{pmatrix} 3 & 4 \end{pmatrix} + 3 \begin{pmatrix} 4 & 5 \end{pmatrix} = \begin{pmatrix} 19 & 25 \end{pmatrix}$$

If A is a matrix, what does:

$$\begin{pmatrix} 4 & 1 & 3 \end{pmatrix} A$$

mean?

the result is four times the first row of A plus the second row, plus three times the third!

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

If A is a matrix, what does:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 1 & 3 \end{pmatrix} A$$

mean?

The first row describes... the first row of the answer, the second describes the second

Interpretation of Matrix Multiplication

If A is a matrix, how do you interpret:

$$A\begin{pmatrix}1\\2\\3\end{pmatrix}?$$

We get the first column of A plus twice the second column, plus three times the column!

$$\begin{pmatrix}1 & 3 & 4\\2 & 4 & 5\end{pmatrix}\begin{pmatrix}1\\2\\3\end{pmatrix} = 1\begin{pmatrix}1\\2\end{pmatrix} + 2\begin{pmatrix}3\\4\end{pmatrix} + 3\begin{pmatrix}4\\5\end{pmatrix} = \begin{pmatrix}19\\25\end{pmatrix}$$

If A is a matrix, how do you interpret:

$$A\begin{pmatrix}4\\1\\3\end{pmatrix}?$$

Four times the first column of A plus the second column, plus three times the column!

If A is a matrix, how do you interpret:

$$A\begin{pmatrix}1 & 4\\2 & 1\\3 & 3\end{pmatrix}?$$

The first column describes the first column, the second describes the second

Formula Definition of Matrix Multiplication

If A is an m by n matrix and B is an n by ℓ matrix we can write:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \vdots \\ a_{m1} & & \cdots & a_{mn} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1\ell} \\ b_{21} & b_{22} & & \\ \vdots & & \ddots & \vdots \\ b_{n1} & & \cdots & b_{n\ell} \end{pmatrix}$$

Then

$$AB = \begin{pmatrix} \sum_{k=1}^m a_{1k} b_{k1} & \sum_{k=1}^m a_{1k} b_{k2} & \cdots & \sum_{k=1}^m a_{1k} b_{k\ell} \\ \sum_{k=1}^m a_{2k} b_{k1} & \sum_{k=1}^m a_{2k} b_{k2} & & \\ \vdots & & \ddots & \vdots \\ \sum_{k=1}^m a_{mk} b_{k1} & & \cdots & \sum_{k=1}^m a_{nk} b_{km} \end{pmatrix}$$

which is summarized by saying that the ij entry of AB is exactly $\displaystyle\sum_{k=1}^{m} a_{ik} b_{kj}$ or if we write

$$AB = C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1\ell} \\ c_{21} & c_{22} & & \\ \vdots & & \ddots & \vdots \\ c_{n1} & & \cdots & c_{n\ell} \end{pmatrix} \quad \text{then have} \quad c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}$$

If you need to prove things about products, sometimes you will need to use this notation!

Properties of Matrix Multiplication

Assume A, B, C, D are matricies and that whenever we write a product/sum it makes sense (ie: the sizes work out).
We know the following things:

- $(A+B)(C+D) = AC + AD + BC + BD$.
- $(AB)C = A(BC)$ so we just write this ABC
- It is rarely true that $AB = BA$, so don't go assuming it is.
- The identity matricies Id_n satisfy $\mathrm{Id}_n A = A$ and $B\mathrm{Id}_m = B$.
- It is possible for $AB = 0_{nm}$ even if neither A nor B is zero.

Each of the above describes a rule about matrix multiplication that you can generally take for granted, but how would you prove one of these things?

Proofs about Matricies

In order to prove things about matricies you often need a lot of notation.
Theorem If we have three matricies

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \vdots \\ a_{m1} & & \cdots & a_{mn} \end{pmatrix} B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1\ell} \\ b_{21} & b_{22} & & \\ \vdots & & \ddots & \vdots \\ b_{n1} & & \cdots & b_{n\ell} \end{pmatrix} C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & c_{22} & & \\ \vdots & & \ddots & \vdots \\ c_{\ell 1} & & \cdots & c_{\ell k} \end{pmatrix}$$

so that the sizes line up, A is m by n, the matrix B is n by ℓ and finally C is ℓ by k, then the two expressions

$$(AB)C \qquad A(BC)$$

are equal.

Proof
The first thing we notice is that in both cases the result will be an m by k matrix (combining m by n with n by ℓ with ℓ by k)
We still need to convince someone they have the same entries... that is, the ij entry of $(AB)C$ is the same as that of $A(BC)$.

Jumping all the way to a final formula is a bit much, we need to explain how we get a formula for each side lets first describe the entries of AB.

Let $D = AB$ and write

$$AB = D = \begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1\ell} \\ d_{21} & d_{22} & & \\ \vdots & & \ddots & \vdots \\ d_{m1} & & \cdots & d_{m\ell} \end{pmatrix}$$

The key formula to note is that

$$d_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{in}b_{nj} = \sum_{x=1}^{n} a_{ix}b_{xj}$$

notice that n, the top summation index is the common dimension between A and B.

Now let us describe $E = BC$,

$$BC = E = \begin{pmatrix} e_{11} & e_{12} & \cdots & e_{1k} \\ e_{21} & e_{22} & & \\ \vdots & & \ddots & \vdots \\ e_{\ell 1} & & \cdots & e_{\ell k} \end{pmatrix}$$

now we have the formula:

$$e_{ij} = b_{i1}c_{1j} + b_{i2}c_{2j} + b_{i3}c_{3j} + \cdots + b_{im}c_{mj} = \sum_{y=1}^{m} b_{iy}c_{yj}$$

Now lets look at $F = (AB)C = DC$ and write

$$(AB)C = DC = F = \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1k} \\ f_{21} & f_{22} & & \\ \vdots & & \ddots & \vdots \\ f_{m1} & & \cdots & f_{mk} \end{pmatrix}$$

The key formula to note is that

$$f_{ij} = d_{i1}c_{1j} + d_{i2}c_{2j} + d_{i3}c_{3j} + \cdots + d_{im}c_{mj} = \sum_{y=1}^{m} d_{iy}c_{yj}$$

notice that m, the top summation index is the common dimension between D and C.
Now let us describe $G = A(BC) = AE$,

$$A(BC) = AE = G = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1k} \\ g_{21} & g_{22} & & \\ \vdots & & \ddots & \vdots \\ g_{m1} & & \cdots & g_{mk} \end{pmatrix}$$

now we have the formula:

$$g_{ij} = a_{i1}e_{1j} + a_{i2}e_{2j} + a_{i3}e_{3j} + \cdots + a_{in}e_{nj} = \sum_{x=1}^{n} a_{ix}e_{xj}$$

notice that n, the top summation index is the common dimension between B and C.

Now we just need to convince ourselves that for all i and j in the appropriate range we have that $f_{ij} = g_{ij}$. We have a formula for each, so we want to compare formulas!

We will do some calculations with these, first we note

$$f_{ij} = \sum_{y=1}^{m} d_{iy} c_{yj} \qquad g_{ij} = \sum_{y=1}^{n} a_{ix} e_{xj}$$

but we have the formulas

$$d_{ij} = \sum_{x=1}^{n} a_{ix} b_{xj} \qquad \Rightarrow \qquad d_{iy} = \sum_{x=1}^{n} a_{ix} b_{xy}$$

and

$$e_{ij} = \sum_{y=1}^{m} b_{iy} c_{yj} \qquad \Rightarrow \qquad e_{xj} = \sum_{y=1}^{m} b_{xy} c_{yj}$$

plugging these into the equations for f_{ij} and g_{ij} we have:

$$f_{ij} = \sum_{y=1}^{m} \left(\sum_{x=1}^{n} a_{ix} b_{xy} \right) c_{yj} \qquad g_{ij} = \sum_{x=1}^{n} a_{ix} \left(\sum_{y=1}^{m} b_{xy} c_{yj} \right)$$

now we simply regroup the summations to get

$$f_{ij} = \sum_{x=1}^{n} \sum_{y=1}^{m} a_{ix} b_{xy} c_{yj} \qquad g_{ij} = \sum_{x=1}^{n} \sum_{y=1}^{m} a_{ix} b_{xy} c_{yj}$$

from which we conclude the equality.

What we just showed is that the ij entry of both matricies

$$(AB)C \qquad A(BC)$$

satisfy the same formula

$$f_{ij} = \sum_{x=1}^{n} \sum_{y=1}^{m} a_{ix} b_{xy} c_{yj} = g_{ij}$$

this means that they are the same, and so

$$(AB)C \qquad A(BC)$$

If you want to convince people that two matricies are the same, you probably want to first convince them they at least have the same size, and then convince them all the entries are the same.
Working with the entries of a matrix will often involve notation.
This is something you will need to do a few times in the course.

Working with notation, and writing proofs using it, is something that takes practice. It is hard in a very different way than other sorts of problems.

Inverse of a matrix

Given a matrix A, we say that a matrix B is the inverse of A if and only if:

$$AB = \mathrm{Id}_n = BA$$

If you interpret a matrix as a description of a systematic way to make modifications matricies, then you can interpret the inverse of a matrix A as:
A systematic way to make modifications to a matrix that will undo what A does.
Such a thing may not even exist
If it does, we say A is invertible.

- to be invertible A **must** be a square.
- there are other conditions... which ones do you know?

Proving things are inverses

Given a matrix A, the best way to prove it is invertible is to find a matrix B so that

$$AB = \mathrm{Id}_n = BA$$

Ideally you can prove this without talking about the entries of the matrix A

If you need to work with matricies explicitly you may be able to this check as follows:

If we have square n by n matricies

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \vdots \\ a_{n1} & & \cdots & a_{nn} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & & \\ \vdots & & \ddots & \vdots \\ b_{n1} & & \cdots & b_{nn} \end{pmatrix}$$

Then we can check if they are inverses by looking at the entries of the products AB (or BA), if they are inverses then by looking at the formula ij entry of AB we get

$$(AB)_{ij} = \sum_{x=1}^{n} a_{ix} b_{xj} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

By looking at BA we get the condition

$$(BA)_{ij} = \sum_{x=1}^{n} b_{ix} a_{xj} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

So to prove A and B are inverses, prove these formulas work out.

Interpretation of Gaussian Elimination

Gaussian elimination is a useful method to solve a system of equations, or find an inverse, but what is it doing?

If I want to perform Gaussian elimination on a matrix A, I do this in a series of steps.

- We start with $A_0 = A$.
- We then perform a series of elementary operations to the matrix, at each stage obtaining $A_i \Rightarrow A_{i+1}$.
- If the ith step is a transformation by an elementary operation E_i Then

$$A_i = E_i A_{i-1} \qquad\qquad A_{i-1} \overset{E_i}{\Rightarrow} A_i$$

The hope is to end with A_n being the identity matrix.
The inverse of A is then ... the combination of all of these elementary steps, that is:

$$A^{-1} = E_n E_{n-1} \cdots E_2 E_1$$

Rather than computing that product all at once at the end, at each stage we also start with B_0 being the identity matrix.

$$B_i = E_i B_{i-1}$$

So that in the end $B_n = E_n E_{n-1} \cdots E_2 E_1 = A^{-1}$.

Interpretation of Gaussian Elimination

If we do this Gaussian elimination on A looks like:

$$\begin{array}{rcl} \big(A & \mid & \mathrm{Id}_n\big) \\ \big(E_1 A & \mid & E_1\big) \\ & \vdots & \\ \big(E_{n-1}E_{n-2}\cdots E_1 A & \mid & E_{n-1}E_{n-2}\cdots E_1\big) \\ \big(R & \mid & E_n E_{n-1}E_{n-2}\cdots E_1\big) \end{array}$$

Where R is in RREF, which will be Id_n if A was invertible.
Notice that this also allows us to write

$$R = (E_n E_{n-1} \cdots E_1)\, A$$

so

$$P = (E_n E_{n-1} \cdots E_1)$$

is a matrix so that PA is upper triangular.

Theorem

Every (square) matrix comes from a product of elementary matricies. (adding a multiple of one row to another, swapping rows, multiplication by a scalar [in this theorem the scalar can be zero, but that is a bad way to solve systems of equations])

Taking Determinants

For a matrix A denote by $\tilde{A}_{k\ell}$ the matrix obtained by deleting the kth row and the ℓth column. Define the determinant of a 0 by 0 matrix to be 1. For an n by n matrix $A = (a_{ij})$ define:

$$\det(A) = \sum_{\ell=1}^{n} (-1)^{\ell+k} a_{k\ell} \det(\tilde{A}_{k\ell}) = \sum_{k=1}^{n} (-1)^{\ell+k} a_{k\ell} \det(\tilde{A}_{k\ell})$$

and it is a strange sort of magic that this doesn't depend on which row or column you expand on.
And it is likewise a fun fact that if we define B by the formula

$$b_{k\ell} = (-1)^{k+\ell} \det(\tilde{A}_{\ell k})$$

then we get

$$AB = \det(A)\mathrm{Id}_n$$

Determinants also compute volumes, and have other strange properties, For some of these, see the handout on Moodle

Proofs about determinants: There are some notes on Moodle showing you some versions of these proofs, such proofs are generally notationally awkward if you don't use advanced ideas beyond what we cover in the course.

Transpose of a matrix

Another thing we can do with matricies, is interchange the role of rows and columns. eg:

$$\begin{pmatrix} 1 & 2 & 7 & 8 \\ 3 & 5 & 6 & 9 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 6 \\ 8 & 9 \end{pmatrix}$$

In many ways, this is just changing how we display our information, however, as you will see on the assignments, this operation actually can be useful to solve problems.

Simple Example
Interpret the vectors

$$\text{2cups of sugar} + \text{2cups of butter} + \text{4eggs} + \text{(1/2)cups of flour} + \text{1cup of cocoa}$$

and

$$\text{3\$/cup sugar} + \text{2\$/cup butter} + \text{1\$/egg} + \text{1\$/cup flour} + \text{1\$/cup of cocoa}$$

as

$$I = \begin{pmatrix} 2 \\ 2 \\ 4 \\ 1/2 \\ 1 \end{pmatrix} \qquad P = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 2 \\ 1 \end{pmatrix} \quad \text{then} \quad I^t P = \big(2(3) + 2(2) + 4(1) + (1/2)(2) + 1(1)\big) = (16)$$

is the cost.

Proofs about Transposes

If you need to prove something involving A^t, the key thing to track notationally is that if we define

$$B = A^t$$

Then the relationship between:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \vdots \\ a_{n1} & & \cdots & a_{nm} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & & \\ \vdots & & \ddots & \vdots \\ b_{n1} & & \cdots & b_{mn} \end{pmatrix}$$

is that

$$b_{xy} = a_{yx}$$

You should expect proofs about Transposes to involve notation.