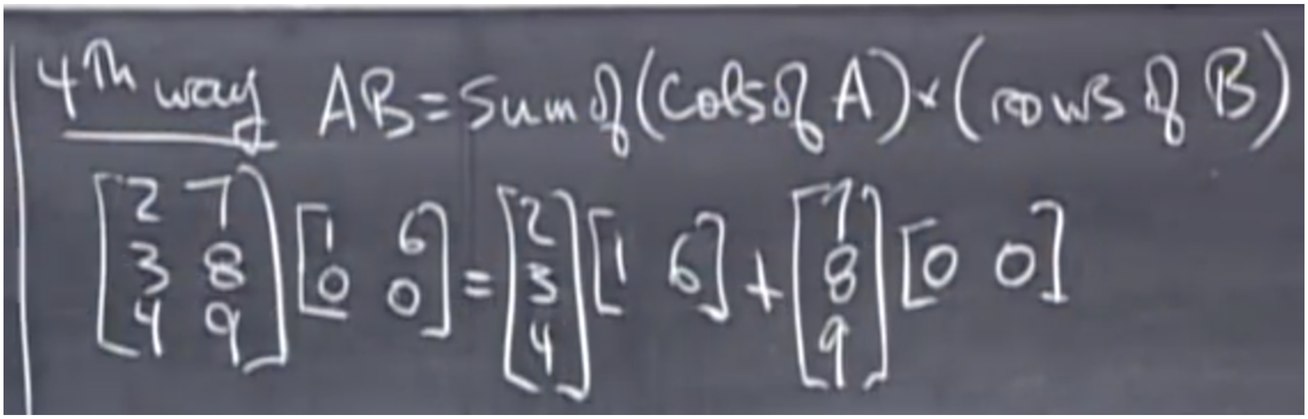
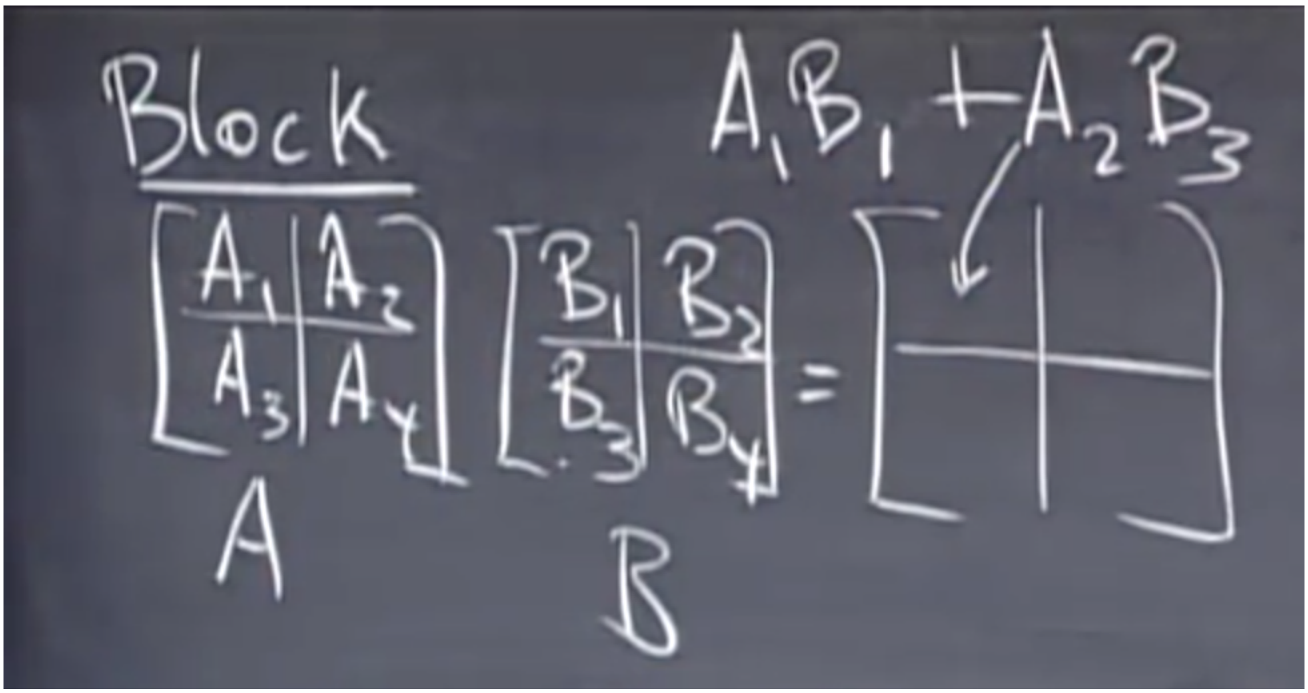
the basics: elimination

- linear algebra is concerned with linear equations, and their solutions. at least that's where things start.
- we saw the definition of linear combination earlier in the differential equations part of our multivariable calc course (it is something where $f(cx) = cf(x)$ and $f(x+y) = f(x) + f(y)$
- we can represent a linear equation in matrix form with the coefficients of each equation in a row as $A\vec{x} = b$ and from this we can graph the two lines on a coordinate axes. This is the row picture. $A\vec{x} = b$ is really just saying $\begin{bmatrix} 2 & 3 \\ 1 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$
- The column picture is a different representation of the same linear system that focuses on the columns: $x\begin{bmatrix} 2 \\ 1 \end{bmatrix} + y\begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$. Note that this picture emphasizes the "linear combination of the columns" way of looking at the system of equations.
- In the column picture, we draw the columns as vectors. The vector on the RHS is the target vector. The question is what should the coefficients of the linear combination be so that the linear combinations of the vectors gives us the target vector.
- Again, note how the row and column pictures are both encoding the same system of equations.
- Given this column picture in terms of linear combinations of vectors, a question that turns out to be important is: given the lhs with coefficients x, y, what are all the achievable rhs vectors? this is important because it tells us for what values of the constants does the equation have a solution?
- Note the row picture asked: where do we intersect? The column picture asks: what is the right values for the coefficients so that the linear combination gives the vector on the rhs? They are different ways of looking at the same problem.
- Similarly we can express equations in more dimensions using either picture.
- Again in 3 dimensions, the row picture involves drawing planes in 3-d and looking for intersection (linear equation in 3-d represents a plane). the row picture involves drawing planes in 3d and this gets hard to visualize.
- the column picture in 3-d involves a linear combination of 3 vectors each of which contain 3 elements. so its something like $xA_1 + yA_2 + zA_3 = \vec{b}$ where A_n is the n^{th} column of the matrix coefficient matrix.
- again, an interesting question is for what right hand sides can we make the linear combination add up. OR do the linear combinations of the columns fill up 3-d space?
- notice that we have shown that when we multiply a matrix by a vector, we get a linear combination of the columns of the matrix.
- in 3-D, if all the column vectors lie in the same plane, then all linear combinations will also lie in that plane and in that case we know that the linear combinations of those columns don't describe 3-d space.
- this kind of "can we do linear combinations of these vectors to achieve any rhs" applies even to higher dimensions → 8-d space, 9-d space etc. in other words "do these 9 vectors in 9-d space fill out all of 9-d space".
- The big takeaway from this lecture is viewing the system of equations (the matrix $A\vec{x} = b$) as a linear combination of columns.
- If you have a 3 x 3 matrix, elimination is the process of reducing the system to an upper triangular matrix, that is one where the elements below the leading diagonal are all 0's. this is basically the same as row reduction that we saw in herb gross' multivariable calc course.

- while doing elimination (row reduction) we arrange the equations so that the one that has a non-zero x coefficient is the first equation, the one that has a non-zero y coefficient is the second equation etc. these non-zero coefficient terms on the diagonal is called the pivot. the pivot is used to reduce the coefficients of the corresponding terms in the below equations to 0's. this gives us the upper triangular matrix(we call this U). Note that we don't accept 0's as pivots. we just interchange until we don't have a 0 there.
- how could elimination fail? how can we fail to come up with 3 non-zero pivots? if we keep row reducing in different combinations and we can't make all non-0 pivots, then we fail → we can't do elimination on this kind of system.
- after we do elimination and assuming its successful, we can plug in values and find the unknowns. this step is called back substitution.
- we have seen in the previous course on multivariable calc the rules of matrix multiplication and that they are multipliable if the number of columns of first is number of rows of second and the resultant has as many rows as first and as many columns as second.
- this process of elimination can be done with matrix notation if we make up some suitable rules:
  - we are allowed to re-write a row as some constant times itself plus/minus another constant times another row and we are allowed to do this how many ever times. (the equivalent of this in the linear equation system is just adding/subtracting both sides of two equations which is of course allowed).
- alternately, we can also express elimination (row reduction) as a matrix multiplication. once we know what additions and subtractions we want, we can make a matrix, pre-multiplying by which achieves the same upper triangular matrix as we got by doing elimination explicitly. this matrix that does this is called the elementary matrix.
- to get an elementary matrix to take another matrix thru a set of row operations, we can easily show that we just need to start with the identity matrix, and do those operations on the identity matrix. Then if we multiply this elementary matrix with the matrix we want those operations to happen on, the multiplication will make it happen. pretty cool!
- in order to make column operations happen (eg. interchange two columns), we start with the identity matrix, do those column operations on it and then multiply the original matrix with this elementary matrix and the resultant matrix will be such that those column operations were done on the original matrix.
- so once we know the row operations involved in elimination, we can express this as a set of elementary matrices; once we do that multiplying these matrices gives us a matrix that has the same effect as the overall elimination operation.
- We call the initial matrix A and the reduced upper triangular matrix U. If we have a matrix E which when applied to another matrix performs some set of operations on it, the matrix that gives us back the original matrix (undoes the effect of multiplying by E) is referred to as E^{-1} (pronounced E inverse).
- There are five ways to look at matrix multiplication:
  - $C_{ij} = \sum_{k=1}^{n} A_{ik} * B_{kj}$
  - the columns of C are a linear combination of the columns of A (with weights from B)
  - the rows of C are a linear combination of the rows of matrix B (with weights from A)
  - C is the sum of (the columns of A * rows of B)



- block multiplication:



- Invertible matrices are such that $AA^{-1} = I$. Left inverses and right inverses exist. Later we will show that for square matrices the left inverse is the same as its right inverse.
- Invertible matrices are also called non-singular. These are the nice ones. The other set is non-invertible (or singular) matrices.
- If a matrix has $A\vec{x} = 0$ with non-0 $\vec{x}$, it must be singular (non-invertible). Proof: Assume that for a non-0 $\vec{x}$, we have $A\vec{x} = 0$ and that A is invertible. If we pre-multiply both sides by A^{-1}, we get $\vec{x}$ (non-zero) on the left side and $\vec{0}$ on the rhs, which is a contradiction.
- Since $Ax = b$ is just a linear combination of the columns of A, if the columns of A have the same direction, then we can't really span the space since the linear combinations of vectors in the same direction are in the same direction as the vectors.
- To find the inverse of A, lets start with $[A|I]$. We do row reduction (row operations) on A and the same operations on I until A goes to I. Whatever the right side matrix becomes (which was initially I) is the elementary matrix which when multiplied with A gives us I. Thus that is A^{-1}. This is called the Gauss-Jordan elimination.

- For a square matrix, proof that right inverse is equal to left inverse. Let the left and right inverses be L and R.

  By definition of inverse, $LA = I \rightarrow LAR = IR \rightarrow LI = IR \rightarrow L = R$.
- Inverse of a product AB is $B^{-1}A^{-1}$. Proof: multiply and verify you get identity.
- transpose is the operation of interchanging rows and columns.
- We can easily show that the inverse of a transpose of a matrix is the transpose of its inverse. That is, if $AA^{-1} = I$ then $A^T(A^{-1})^T = I$.
- If we have a nice matrix A which can be row reduced *without row exchanges,* we can come up with a series of elementary matrices that reduce that matrix to the upper triangular form U.
- If we have a nice matrix (can be row reduced without row exchanges), the product of the this elementary matrices is $E_{ij}..E_{pq}$. We can easily show that the inverse of this product, $E_{pq}^{-1}...E_{ij}^{-1}$, is a lower triangular matrix. So we have $EA = U \rightarrow A = E^{-1}U \rightarrow A = LU$.
- Further we can break out U into just a matrix with pivots on the diagonals and all others zeros (called the diagonal matrix) and another matrix U using the idea that a matrix multiplication is just a linear combination of rows of the matrix in the right side of the multiplication. With this breakup, we get $A = LDU$. Again, all this is only true for nice matrices.
- If we have a nice matrix of size n x n, we can easily show that row reduction is $O(n^3)$
- A permutation matrix is an identity matrix with interchanged rows. It can be used as a pre-multiplier to interchange rows of other matrices. Or as a post multiplier to interchange columns of other matrices. The number of permutation matrices for an n x n matrix is $n!$ (obviously).
- For a identity matrix of size n x n, the set of permutation matrices are closed with respect to multiplication and inverses. the inverse of a permutation matrix is just its transpose, so $PP^T = I$. (think about why this is true). Hence $P^T = P^{-1}$.
- If a matrix isn't "nice" (reducible without row permutations), but is invertible, we can write a $PA = LU$ where P is a permutation matrix that makes A nice (identity matrix with re-ordered rows).
- definition of transpose: the elements are such that $A_{ij}^T = A_{ji}$. Matrices where $A = A^T$ are called symmetric matrices. These are reflections on their leading diagonal.
- We can easily show that for any two matrices, R^TR is always symmetric. this is even true for rectangular matrices, not just square matrices.
- We can also easily show that $(AB)^T = B^TA^T$.