These are brief notes for the lecture on Wednesday September 15, 2010: they are not complete, but they are a guide to what I want to say today. They are not guaranteed to be correct.

2.4. Partitioned Matrices

Sometimes it is helpful to group certain rows and columns of a matrix together, and to regard the entries as one object: for example, the 6×6 matrix

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

looks rather like a diagonal matrix: except that down the diagonal we have 2×2 matrices instead of numbers. We can view this as

$$\begin{pmatrix} J & 0 & 0 \\ 0 & J & 0 \\ 0 & 0 & J \end{pmatrix}$$

in which each of the entries J and 0 ("blocks") is a 2×2 matrix. Sometimes this is a very useful view to take.

Things can get more complicated: the pattern of blocks might be such that different blocks have different sizes: However, all the blocks in a row of blocks have to have the same number of rows, and all the blocks in a column of blocks have to have the same number of columns.

Example:

We can specify the sizes of the blocks in a partitioned $m \times n$ matrix by $(m_1, m_2, \ldots, m_k)$ and $n_1, n_2, \ldots, n_l)$, where the m_i are positive integers summing to m and the n_j are positive integers summing to n. Then the partitioned matrix consists of k rows of blocks and l columns of blocks.

If A and B are two partitioned matrices with the same block sizes (i.e. the list of m_i and n_j values are the same) then we can add A and B by adding each of their blocks. To multiply a partitioned matrix by a scalar, just multiply each block by the scalar.

Matrix multiplication is (surprise!) more complicated! Partitioned matrices A and B can be multiplied together (in a manner respecting the partition) if each block of A can be multiplied into the corresponding block of B: that is, if the partitioning of the columns of A is the same as the partitioning of the rows of B.

In this case, things work out just as we would expect:

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}$$

A special case of this is when the blocks of A are just the columns of A, the blocks of B are just the rows of B, and, of course, the number of columns of A is equal to the number of rows of B. Then

$$A = [\mathrm{col}_1(A), \mathrm{col}_2(A), \ldots, \mathrm{col}_n(A)] \qquad \text{and} \qquad B = \begin{pmatrix} \mathrm{row}_1(B) \\ \mathrm{row}_2(B) \\ \vdots \\ \mathrm{row}_n(B) \end{pmatrix}$$

If A is $m \times n$ and B is $n \times p$ then each $\mathrm{row}_i(A)\mathrm{col}_j(B)$ is a $m \times p$ matrix, and

$$AB = \mathrm{col}_1(A)\mathrm{row}_1(B) + \mathrm{col}_2(A)\mathrm{row}_2(B) + \ldots \mathrm{col}_n(A)\mathrm{row}_n(B)$$

Example: Compute $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 5 \\ 6 & 5 \\ 4 & 3 \end{pmatrix}$ in this manner.

Inverses of block matrices are rather more complicated. We won't touch much on them here.

2.5. Matrix Factorizations

It is often the case that we need to repeatedly solve matrix equations $A\underline{x} = \underline{b}_1$, $A\underline{x} = \underline{b}_2, \ldots$, $A\underline{x} = \underline{b}_p$ (solving for different values of $\underline{x}$ each time, of course). If A is invertible, of course, it would be easy to compute A^{-1} and just compute $A^{-1}\underline{b}_1$, etc. However, there is another efficient method, one which is actually used in practice. This is the so-called LU decomposition or factorization of A.

Suppose that A is a $m \times n$ matrix which can be row-reduced to echelon form (*not* reduced row echelon form!) without switching rows. Then if the elementary row operations have matrices E_1, E_2, $\ldots E_k$, then we get $U = E_k E_{k-1} \ldots E_2 E_1 A$ is in echelon form. Now $E_k E_{k-1} \ldots E_2 E_1$ is the product of lower triangular $m \times m$ matrices (that is, matrices for which the entries above the diagonal are 0) and each of them has 1's down the diagonal (since for reduction to echelon form we don't divide rows by a scalar to make the leading entry equal to 1).

Hence A can be written as LU, where $L = E_1^{-1} E_2^{-1} \ldots E_k$ is lower triangular, and has 1's down the diagonal (it is a *unit lower triangular matrix*) and U is in echelon form.

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{pmatrix} \begin{pmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Here $*$ refers to any real number, and $\blacksquare$ marks the pivots in U.

To solve the equation $A\underline{x} = \underline{b}$ we see that it is the same as solving $LU\underline{x} = \underline{b}$. If we first solve the equation $L\underline{y} = \underline{b}$ then the equation $U\underline{x} = \underline{y}$ we will solve the system.

But $L\underline{y} = \underline{b}$ is easy to solve, since L is lower triangular. And since U is upper triangular, the equation $U\underline{x} = \underline{y}$ is also easy to solve.

2.6. Subspaces of $\mathbb{R}^n$

In Chapter 4 we will meet general vector spaces, and study subspaces of them: in order to make the ideas we meet there more concrete, we'll briefly discuss subspaces and dimension of subspaces here.

A *subspace* of $\mathbb{R}^n$ is a set of vectors in $\mathbb{R}^n$ which looks like a copy of $\mathbb{R}^m$ in its own right: for example, if we are in $\mathbb{R}^3$, any plane through the origin looks a lot like $\mathbb{R}^2$: it is "2-dimensional", the sum of any two vectors in the plane is in the plane, and any scalar multiple of a vector in the plane is in the plane. Similarly, any line through the origin looks a lot like $\mathbb{R} = \mathbb{R}^1$. How can we formalize this notion, and get our hands on what "dimension" means?

DEFINITION. *A* subspace *of* $\mathbb{R}^n$ *is a set* H *of vectors in* $\mathbb{R}^n$ *with the following three properties:*

(1) *The zero vector* $\underline{0}$ *from* $\mathbb{R}^n$ *is in* H*.*

(2) *For every* $\underline{u}, \underline{v} \in H$*,* $\underline{u} + \underline{v} \in H$*.*

(3) *For each* $\underline{u} \in H$ *and* $c \in \mathbb{R}$*,* $c\underline{u} \in H$*.*

Example 1: If $\underline{v}_1$ and $\underline{v}_2$ are in $\mathbb{R}^n$, then $\mathrm{Span}(\underline{v}_1, \underline{v}_2)$ is a subspace of $\mathbb{R}^n$. To check this, note first that $\underline{0} = 0\underline{v}_1 + 0\underline{v}_2$.
Now check it is closed under addition:

Now check it is closed under scalar multiplication:

Example 2: A line *not* through the origin is not a subspace: it doesn't contain $\underline{0}$. Also, it is not closed under scalar multiplication or addition.

Example 3: For $\underline{v}_1, \underline{v}_2, \dots \underline{v}_p \in \mathbb{R}^n$, $\mathrm{Span}(\underline{v}_1, \underline{v}_2, \dots \underline{v}_p)$ is a subspace of $\mathbb{R}^n$. We will refer to this as *the subspace spanned* (or *generated*) by $\underline{v}_1, \underline{v}_2, \dots \underline{v}_p$.

An important special case of this example is the following:

DEFINITION. *The* column space *of a matrix A is the set Col(A) of all linear combinations of the columns of A.*

That is, $\mathrm{Col}(A)$ is the span of the columns of A.

The other common way in which subspaces arise is as the solution set to a homogeneous system of equations:

DEFINITION. *The* null space *of a matrix A is the set Nul(A) of all solutions to the homogeneous equation $A\underline{x} = \underline{0}$.*

THEOREM 12. *The null space of a $m \times n$ matrix A is a subspace of $\mathbb{R}^n$.*