Math 408 Section A WINTER 2004

LINEAR ALGEBRA REVIEW

In this course the notion of **linearity** plays a central role. Many of the theoretical aspects of this course are based on properties of systems of linear equations and inequalities in $I\!R^n$. For this reason the course prerequisite MATH 308 should be taken very seriously. This guide has been prepared to facilitate your review of the relevant material on linear systems of equations. Reviewing this material will prepare you for the quizzes.

Basic Review:

You should be able to answer the following questions:

Answer the following 6 questions for each of the four cases $n = 1$, $n = 2$, $n = 3$, and $n \in \{1, 2, 3, 4, 5, 6, \ldots\}$ beginning with the case $n = 1$.

1. Why is $I\!R^n$ called a vector space?
2. What is the *dot product* (or inner product) on $I\!R^n$?
3. What is the angle between any two vectors in $I\!R^n$? (For this one simply applies the definition $a^T b = \|a\|\|b\| \cos\theta$.)
4. When are two vectors in $I\!R^n$ said to be *orthogonal*?
5. Describe all linear mappings from $I\!R^n$ to $I\!R$.
6. What is the data structure that one usually associates with a linear transformation from $I\!R^n$ to $I\!R^m$?

   You should also be able to answer the following more specific questions:

7. Given any two points in $I\!R^2$ determine (a) an equation for the line passing through these points, and (b) a vector normal to the line determined by these two points.
8. Given any three non–collinear points in $I\!R^3$ determine (a) an equation for the plane passing through these three points, and (b) a vector normal to this plane.
9. Let $a = (a_1, a_2, \ldots, a_n)^T \in I\!R^n$ and $\alpha \in I\!R$. For $n = 1, 2, 3$ and $n > 3$ describe the set

   $$\{x \in I\!R^n \ : \ a^T x = \alpha\}$$

   when (a) $\alpha = 0$, and (b) $\alpha \neq 0$.
10. Let $a_{1\cdot} = (a_{11}, a_{12}, \ldots, a_{1n})^T \in I\!R^n$, $a_{2\cdot} = (a_{21}, a_{22}, \ldots, a_{2n})^T \in I\!R^n$, and $\alpha_1, \alpha_2 in I\!R$. For $n = 1, 2, 3$ and $n > 3$ describe the set

    $$\{x \in I\!R^n \ : \ a_{1\cdot}{}^T x = \alpha_1, \ a_{2\cdot}{}^T x = \alpha_2\}.$$
11. For $i = 1, 2, 3, \ldots, m$, let $a_{i\cdot} = (a_{i1}, a_{i2}, \ldots, a_{in})^T \; in I\!R^n$ and $\alpha_i \in I\!R$.

(a) For $n = 1, 2, 3$ and $n > 3$ with $m = n$ describe the set

$$S = \{x \in I\!R^n \ : \ a_{i\cdot}{}^T x = \alpha_i, \ i = 1, 2, \ldots, m\}.$$

(b) How does this description change when (a) $m < n$, (b) $m > n$?

(c) Can you express the set S in matrix notation?

(d) What does it mean to say that the underlying matrix has (a) full column rank, (b) full row rank, (c) full rank?

12. What are the three elementary row operations, i.e. the three operations associated with row reduction?

13. Given an $n \times n$ real matrix A, what does it mean to say that A is invertible (or, equivalently, non–singular)?

14. Let A be an $m \times n$ real matrix, B be a $k \times n$ real matrix, and let $a \in I\!R^m$ and $b \in I\!R^k$. We say that the two systems of equations

$$Ax = a \qquad \text{and} \qquad Bx = b$$

are equivalent if they have identical solution sets.

(a) Show that if any one of the three elementary row operations is applied to the system $Ax = a$, then one obtains an equivalent linear system.

(b) If C is a non–singular $m \times m$ real matrix, show that the system $Ax = a$ is equivalent to the system $CAx = Ca$.

(c) Is it possible for the systems $Ax = a$ and $Bx = b$ to be equivalent when $n \neq m$? If it is possible, then provide an example to illustrate this possi bility.

Advanced Review:

These notes elaborate on a number of important topics studied in Math 308 (Introduction to Linear Algebra). One of the goals of these notes is to discuss the *ideas* in the course without getting too bogged down in the notational details. However, at certain points this will become unavoidable. Hopefully, this approach will help you see the big picture a little better. But be forewarned, these notes are not an easy read! Reading these notes is an exercise in itself since they require a familiarity with many of the terms and ideas defined throughout the course. The notes exercise your knowledge of this material and are intended to solidify this knowledge.

Subspaces:

Recall that a subset W of $I\!R^n$ is a subspace if and only if it satisfies the following three conditions:

1. The origin is an element of W.
2. The set W is closed with respect to addition, i.e. if $u \in W$ and $v \in W$, then $u + v \in W$.

3. The set W is closed with respect to scalar multiplication, i.e. if $\alpha \in I\!R$ and $u \in W$, then $\alpha u \in W$.

For example, by using this fact it is easily shown that given any set S in $I\!R^n$ we have that the set

$$S^{\perp} = \{v \,:\, w^T v = 0 \text{ for all } w \in S\}$$

is a subspace. If S is itself a subspace, then $S^{\perp}$ is called the subspace orthogonal (or perpendicular) to the subspace S. Moreover, in this case we have $S = (S^{\perp})^{\perp}$ (more generally, $(S^{\perp})^{\perp} = \text{Span}\,(S)$). If S is a subspace, it can be shown that

$$n = \dim(S) + \dim(S^{\perp})\ . \tag{1}$$

Every subspace has what I will call *internal* and *external* representations.

An *internal* representation is any representation of the subspace as a linear span of a finite set of vectors. The representation is said to be internal since the spanning vectors lie within the subspace. If the set of spanning vectors happens to be linearly independent, then it is called a *basis* of the subspace. It is known that every basis of a subspace has the same number of vectors in it. This number is called the *dimension* of the subspace.

Internal representations of a subspace can be interpreted with the aid of our notion of matrix–vector multiplication. Recall that matrix–vector multiplication can be viewed as taking a linear combination of the columns of the matrix. Thus, if a subspace is known to be the linear span of a finite collection of vectors, then this subspace is the same as the range of the matrix formed by taking the columns of the matrix to be the vectors that span the subspace.

An *external* representation of a subspace is any representation of the subspace as the intersection of a finite number of subspaces of the form $\{x \in I\!R^n \,:\, v_i^T x = 0\}$ for some nonzero v_i in $I\!R^n$ for $i = 1, 2, \ldots, k$. The representation is said to be *external* since the vectors v_i clearly cannot belong to the subspace. Another way to view an external representation of the subspace is that the vectors $\{v_1, v_2, \ldots, v_k\}$ form a spanning set for the subspace *orthogonal* (or perpendicular) to the subspace we are interested in. If the dimension of the subspace is p, then the dimension of the orthogonal subspace is $n - p$. Thus, in particular, $k \geq n - p$.

External representations of a subspace can also be interpreted with the aid of our notion of matrix–vector multiplication. Recall that matrix–vector multiplication can also be interpreted as taking the dot product with each row of the matrix. Thus, a vector is in the null space of a matrix if it is orthogonal to every row of that matrix. Consequently, a subspace externally represented by the vectors $\{v_1, v_2, \ldots, v_k\}$ is the same as the null space of the matrix formed by letting its rows be the vectors v_i in $I\!R^n$ for $i = 1, 2, \ldots, k$.

It is important to remember that every subspace has both internal and external representations (indeed, infinitely many of them). Equivalently, every subspace can be represented either as the range of some matrix or as the null space of some matrix (indeed, there are infinitely many such matrices). Some of the computational techniques we have learned in this course deal with passing between such representations and obtaining minimal representations, that is internal and external representations having the fewest number of elements.

Subspaces Associated with Matrices:

We have seen that matrices and subspaces are intimately connected to each other. In this regard, every matrix has associated with it four fundamental subspaces; its range and null space and the range and null space of its transpose. Due to the interpretations of these subspaces given above, these subspaces have significant connections to each other. Given a matrix $A \in I\!R^{m \times n}$ we have

$$\begin{aligned}
\mathrm{Null}\,(A) &= (\mathrm{Ran}\,(A^T))^{\perp} && (2)\\
\mathrm{Null}\,(A^T) &= (\mathrm{Ran}\,(A))^{\perp} && (3)\\
\dim(\mathrm{Ran}\,(A)) &= \dim(\mathrm{Ran}\,(A^T)) && (4)\\
n &= \dim(\mathrm{Ran}\,(A)) + \dim(\mathrm{Null}\,(A)) && (5)\\
m &= \dim(\mathrm{Ran}\,(A^T)) + \dim(\mathrm{Null}\,(A^T))\ . && (6)
\end{aligned}$$

The dimension of the range of a matrix is called the *rank* of the matrix (denoted $\mathrm{rank}\,(A)$) and the dimension of the null space of a matrix is called the *nullity* of the matrix (denoted $\mathrm{nullity}\,(A)$). Thus, from equation (4) , we have $\mathrm{rank}\,(A) = \mathrm{rank}\,(A^T)$. However, equations (5) and (6) indicate that $\mathrm{nullity}\,(A) = \mathrm{nullity}\,(A^T)$ if and only if $m = n$. Thus, in the non-square case, it will never be the case that $\mathrm{nullity}\,(A)$ equals $\mathrm{nullity}\,(A^T)$. It should also be noted that equations (2) and (3) imply that equations (5) and (6) are special cases of equation (1).

Echelon Form:

The key computational tool that we have employed in this class is Gaussian elimination. Gaussian elimination was introduced for the purpose of transforming a matrix into *echelon form*. That is, Gaussian elimination was introduced as a tool for transforming a matrix into another matrix having the property that the first nonzero entry in the $(i+1)$st row lies to the right of the first nonzero entry in the ith row. Let us review this process in the light of our knowledge of subspaces.

Recall that in Gaussian elimination one employs the three elementary row operations to put a matrix in upper triangular form (or echelon form). These operations can be viewed as operations on the vectors that form the rows of the matrix. That is, they are operations in the row space of the matrix (recall that the row space of a matrix is the linear span of the rows of the matrix, or equivalently, it is the subspace formed by taking all possible linear combinations of the rows of the matrix, or equivalently, it is the range of the transpose of the matrix). In class it was shown that the three elementary row operations do not change the row space of the matrix. Let us review this fact by considering the row operations one at a time.

<u>Row Interchange:</u> In this operation we simply change the order in which we write the rows. Clearly this does not change the row space of the matrix.

<u>Multiply a Row by a Nonzero Scalar:</u> Again the row space remains unchanged since the row space contains all linear combinations of the rows and scalar multiplication is simply a special case of this.

<u>Replace a Row by the Sum of Itself and a Multiple of Another Row:</u> This just represents a linear combination of two rows. Thus the new row remains in the row space.

The essential feature of the three elementary row operations is that *the new collection of rows have the same linear span as the rows you started with.* In order to see this we just need to show that the only row that was changed can be obtained as some linear combination of the new rows. This is obvious for the first two elementary row operations. For the third it is equally obvious since the old row is simply the new row minus what was added to it, namely a scalar multiple of one of the unchanged rows. Thus, the elementary row operations do not change the row space of the matrix. In particular, we have that two matrices are *row equivalent* if and only if they have the same row space.

Now to say that we have put a matrix into echelon form implies that we have made as many rows of the matrix zero as we possibly can. Thus, none of the remaining nonzero rows can be represented as a linear combination of the other remaining nonzero rows. That is, the remaining nonzero rows are linearly independent and have the same span as the row space of the original matrix, or equivalently, the nonzero rows form a basis for the row space of the original matrix. This is a remarkable fact! Our primary tool for solving equations is also an efficient way to obtain a basis for a subspace from a spanning set for that subspace!

Echelon form is indeed a very powerful tool. We have used it over and over again in this course to solve a variety of problems. Let us now take a moment to review the types of problems we have used it to solve.

<u>Equation Solving:</u> Given a system of linear equations, $Ax = b$, where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, we can characterize the solution set to this system by reducing the associated augmented matrix $[A|b]$ to echelon form.

<u>Passing between internal and external representations of a subspace:</u>

<u>Internal to external:</u> This is the same as expressing the range of one matrix as the null space of another. Let A be the matrix in question and assume y is in the range of A. Since $y \in \mathrm{Ran}\,(A)$, the system $Ax = y$ is consistent. We then formally solve this system by reducing the augmented matrix $[A|y]$ to echelon form. Once in echelon form some of the rows will be all zero except for the right–hand side which will be a algebraic expression in the components of y. The consistence of the system implies that each of these expressions must be zero. By writing these expressions in matrix form and setting them equal to zero we get a matrix equation of the form $By = 0$ for some matrix B. That is, we have expressed the range of A as the null space of B.

<u>External to internal:</u> This is the same as expressing the null space of one matrix as the range of another. Again, letting A be the matrix in question, we reduce A to *reduced echelon form.* Typically, this process terminates with a row equivalent matrix having the following block structure:

$$C = \begin{bmatrix} I & B \\ 0 & 0 \end{bmatrix} .$$

Then the columns of the matrix $D = \begin{bmatrix} -B \\ I \end{bmatrix}$ form a basis for the null space of A. In order to see this recall that the null space of A is the subspace orthogonal to the row space of A, that is, $\mathrm{Ran}\,(A^T) = \mathrm{Null}\,(A)^{\perp}$. Since C is row equivalent to A, the null space of A is the subspace orthogonal to the row space of C. Next observe that $CD = 0$, that is, the columns of D are orthogonal to the rows of C. Since the dimension counts work out, the columns of D must be a basis for the null space of C.

Computing a basis for the span of a finite collection of vectors: This has already been discussed above as one of the primary consequences of the echelon form. Just write the vectors as the row vectors of a matrix and then reduce this matrix to echelon form. The rows of the reduced matrix will be the desired basis.

Computing the rank and nullity of a matrix: In this regard, recall the following fundamental theorem:

THEOREM (The Fundamental Theorem on Rank and Nullity)
Let A be an $m \times n$ real matrix and suppose that A is row equivalent to the matrix B in echelon form. If r is the number of non–zero rows in B, then r is the rank of A and $n - r$ is the nullity of A. Thus, in particular, we have

$$\mathrm{rank}\,(A) + \mathrm{nullity}\,(A) = n\ .$$

Computing bases for the four fundamental subspaces associated with a matrix:
Let $A \in \mathbb{R}^{m\times n}$. An efficient procedure for computing bases for the four fundamental subspaces associated with A begins by reducing the augmented system

$$[A\,|\,I] \tag{7}$$

to echelon form. Recall that this is equivalent to multiplying the augmented system on the left by some nonsingular matrix M, yielding

$$[MA\,|\,M]\ . \tag{8}$$

Decomposing this augmented matrix conformally with respect to the nonzero and zero rows of MA yields the block matrix

$$\left[\begin{array}{c|c} T_1 & T_2 \\ 0 & T_3 \end{array}\right], \tag{9}$$

where the matrix T_1 is in echelon form and has no zero rows and $M = \begin{bmatrix} T_2 \\ T_3 \end{bmatrix}$. Then, the rows of the matrix T_1 form a basis for $\mathrm{Ran}\,(A^T)$ and the rows of the matrix T_3 form a basis for $\mathrm{Null}\,(A^T)$. If no row interchanges were required for the reduction to echelon form, then the matrix T_3 has the form

$$T_3 = [T_{31}\ \ I]\ . \tag{10}$$

If this is the case, then the columns of the matrix

$$\begin{bmatrix} I \\ -T_{31} \end{bmatrix} \tag{11}$$

form a basis for Ran (A). Finally, to get a basis for Null (A), we again use Gaussian elimination to transform the matrix T_1 to reduced echelon form, yielding a matrix which typically has block structure

$$[I \; T_{12}] \; . \tag{12}$$

In this case, the columns of the matrix

$$\begin{bmatrix} -T_{12} \\ I \end{bmatrix} \tag{13}$$

form a basis for Null (A).

We now illustrate this process on the matrix

$$\begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 0 \\ -1 & 1 & -4 \\ 3 & 5 & 4 \end{bmatrix} .$$

In this case, the augmented system (7) has the form

$$\left[\begin{array}{rrr|rrrr} 1 & 2 & 1 & 1 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & -4 & 0 & 0 & 1 & 0 \\ 3 & 5 & 4 & 0 & 0 & 0 & 1 \end{array}\right] .$$

After reduction to echelon form, we obtain the matrix

$$\left[\begin{array}{rrr|rrrr} 1 & 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 4 & -3 & 1 & 0 \\ 0 & 0 & 0 & -4 & 1 & 0 & 1 \end{array}\right] .$$

In this case, the matrices T_1, T_2, T_3 and T_{31} appearing in (9) and (10) are

$$T_1 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}, \quad T_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{bmatrix}, \quad T_3 = \begin{bmatrix} 4 & -3 & 1 & 0 \\ -4 & 1 & 0 & 1 \end{bmatrix},$$

and $T_{13} = \begin{bmatrix} 4 & -3 \\ -4 & 1 \end{bmatrix}$. Therefore, as indicated above,

$$\text{Ran}\,(A^T) = \text{Span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}, \quad \text{Null}\,(A^T) = \text{Span}\left\{ \begin{pmatrix} 4 \\ -3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\},$$

and

$$\text{Ran}\,(A) = \text{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ -4 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 3 \\ -1 \end{pmatrix} \right\} .$$

Finally, transforming T_1 to reduced echelon form yields the matrix

$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \end{bmatrix}.$$

Therefore, the matrix T_{12} appearing in (12) is $T_{12} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$, and so

$$\text{Null}\,(A) = \text{Span}\left\{ \begin{pmatrix} -3 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Recapping, we see that beginning with the augmented matrix (7), one can reduce to echelon form to obtain the matrix (9). At this point one can immediately read off bases for both $\text{Ran}\,(A^T)$ and $\text{Null}\,(A^T)$. This leads to an alternative method for computing bases for the subspaces $\text{Ran}\,(A)$ and $\text{Null}\,(A)$. That is to simply repeat the process described above with A replaced by A^T. In this approach one begins with the augmented matrix

$$\left[A^T \mid I\right]$$

and reduces to the echelon form

$$\left[\begin{array}{c|c} S_1 & S_2 \\ 0 & S_3 \end{array}\right], \tag{14}$$

where S_1 is in echelon form and has no zero rows. Then the rows of S_1 form a basis for $\text{Ran}\,(A)$, while the rows of S_3 form a basis for $\text{Null}\,(A)$. Applying this approach to the example given above, we begin by reducing the augmented matrix

$$\left[\begin{array}{cccc|ccc} 1 & 1 & -1 & 3 & 1 & 0 & 0 \\ 2 & 3 & 1 & 5 & 0 & 1 & 0 \\ 1 & 0 & -4 & 4 & 0 & 0 & 1 \end{array}\right]$$

to the echelon form

$$\left[\begin{array}{cccc|ccc} 1 & 1 & -1 & 3 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & -3 & 1 & 1 \end{array}\right].$$

Consequently,

$$S_1 = \begin{bmatrix} 1 & 1 & -1 & 3 \\ 0 & 1 & 3 & -1 \end{bmatrix} \quad \text{and} \quad S_3 = \begin{bmatrix} -3 & 1 & 1 \end{bmatrix}.$$

Therefore,

$$\text{Ran}\,(A) = \text{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 3 \\ -1 \end{pmatrix} \right\} \quad \text{and} \quad \text{Null}\,(A) = \text{Span}\left\{ \begin{pmatrix} -3 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Eigenvalues and Eigenvectors:

Eigenvalues and eigenvectors arise in the context of *square* matrices. Thus, all matrices discussed in this section are assumed to be square. Recall that a scalar λ is called an eigenvalue of the matrix A if there is a nonzero vector x, called an eigenvector, such that $Ax = \lambda x$. One of the beautiful features of eigenvalues and eigenvectors is that they allow us to write matrix–vector multiplication as scalar–vector multiplication.

We now are confronted with two important question:

1. How can we compute the eigenvalues of a matrix?
2. Having an eigenvalue for a matrix, how can we characterize its associated eigenvectors?

The key to answering both of these questions is to observe that λ is an eigenvalue of A with associated eigenvector x if and only if $x \in \text{Null}\,(A - \lambda I)$. Indeed, $x \in \text{Null}\,(A - \lambda I)$ if and only if $(A - \lambda I)x = 0$, or equivalently, $Ax = \lambda x$. The subspace $\text{Null}\,(A - \lambda I)$ is called the eigenspace associated with the eigenvalue λ. Thus, in particular, λ is an eigenvalue of A if and only if the matrix $A - \lambda I$ is singular. Therefore, we can answer the first question with aid of the determinant since the determinant of a square matrix is zero if and only if the matrix is singular. For this purpose, we define the characteristic polynomial for a square matrix A to be the polynomial given by the determinant

$$p(t) = \det\,(A - tI)\ .$$

Then the roots of the characteristic polynomial $p(t)$ coincide precisely with the eigenvalues of the matrix A.

Now having an eigenvalue λ of the matrix A, the set of all eigenvectors associated with λ coincides with the nonzero vectors of the subspace $\text{Null}\,(A - \lambda I)$. The algebraic multiplicity of an eigenvalue is the multiplicity of the eigenvalue as a root of the characteristic polynomial. The geometric multiplicity of an eigenvalue is the dimension of the eigenspace $\text{Null}\,(A - \lambda I)$. We always have that

$$\text{geometric multiplicity} \leq \text{algebraic multiplicity}.$$

An eigenvalue is said to be defective if its geometric multiplicity does not equal its algebraic multiplicity. A matrix is said to be defective if one or more of its eigenvalues is defective.

Further Notes on Matrix Multiplication:

Given two matrices $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{s\times t}$ the matrix product $C = AB$ is defined if and only if $m = s$ in which case C is in $\mathbb{R}^{n\times t}$ with

$$C_{ij} = \sum_{k=1}^{m} A_{ik}B_{kj} \text{ for } i = 1, \ldots, n \text{ and } j = 1, \ldots, t.$$

Another way to view this product is as follows. Let $A_{i\cdot}$ and $A_{\cdot j}$ denote the ith row and jth column of the matrix A, respectively. Then

$$C_{ij} = A_{i\cdot} \bullet B_{\cdot j} \text{ for } i = 1, \ldots, n \text{ and } j = 1, \ldots, t,$$

where $\bullet$ denotes the vector dot product.

Building on this notation, we can define the matrix product of *block structured* matrices. Consider the matrices $M \in \mathbb{R}^{n\times m}$ and $T \in \mathbb{R}^{m\times k}$ given by

$$M = \begin{bmatrix} A_{n_1\times m_1} & B_{n_1\times m_2} \\ C_{n_2\times m_1} & D_{n_2\times m_2} \end{bmatrix}$$

and

$$T = \begin{bmatrix} E_{m_1\times k_1} & F_{m_1\times k_2} & G_{m_1\times k_3} \\ H_{m_2\times k_1} & J_{m_2\times k_2} & K_{m_2\times k_3} \end{bmatrix},$$

where $n = n_1 + n_2$, $m = m_1 + m_2$, and $k = k_1 + k_2 + k_3$. The block structures for the matrices M and T are said to be *conformal* with respect to matrix multiplication since

$$MT = \begin{bmatrix} AE + BH & AF + BJ & AG + BK \\ CE + DH & CF + DJ & CG + DK \end{bmatrix}.$$

Similarly, one can conformally block structure matrices with respect to matrix addition (how is this done?).

Understanding how to manipulate matrices with respect to their block structure is an extremely powerful tool in numerical linear algebra. We give a simple example of the power of this technique by describing Gaussian elimination as a series of matrix multiplies using the underlying block structure of the intermediary resultants.

Given a vector

$$v_1 = \begin{bmatrix} \alpha_1 \\ a_1 \end{bmatrix}$$

in $\mathbb{R}^n$ with $\alpha_1 \in \mathbb{R}$ not equal to zero and $a_1 \in \mathbb{R}^{n-1}$, we define the matrix

$$G_1 = \begin{bmatrix} \alpha_1^{-1} & 0 \\ -\alpha_1^{-1}a_1 & I_{(n-1)\times(n-1)} \end{bmatrix},$$

where $I_{(n-1)\times(n-1)}$ is the $(n-1)\times(n-1)$ identity matrix, as its associated Gaussian elimination matrix. Using the technique of block structured matrix multiplication it is easy to see that

$$G_1 v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Now consider the matrix

$$A_1 = \begin{bmatrix} \alpha_1 & b_1^T \\ a_1 & B_1 \end{bmatrix},$$

where $b_1 \in \mathbb{R}^{m-1}$ and $B_1 \in \mathbb{R}^{(n-1)\times(m-1)}$ so that $A_1 \in \mathbb{R}^{n\times m}$ (assume that $m \geq n$). We have

$$G_1 A_1 = \begin{bmatrix} 1 & \alpha_1^{-1}b_1^T \\ 0 & B_1 - \alpha_1^{-1}a_1 b_1^T \end{bmatrix}.$$

Define $A_2 = B_1 - \alpha_1^{-1} a_1 b_1^T$ and conformally decompose A_2 as

$$A_2 = \begin{bmatrix} \alpha_2 & b_2^T \\ a_2 & B_2 \end{bmatrix},$$

where $\alpha_2 \in I\!R,\ a_2 \in I\!R^{(n-2)},\ b_2 \in I\!R^{m-2}$ and $B_2 \in I\!R^{(n-2)\times(m-2)}$. Assume that $\alpha_2 \neq 0$ and let

$$G_2 = \begin{bmatrix} 1 & 0 \\ 0 & \hat{G}_2 \end{bmatrix},$$

where $\hat{G}_2$ is defined to be the Gaussian elimination matrix associated with the vector

$$v_2 = \begin{bmatrix} \alpha_2 \\ a_2 \end{bmatrix}.$$

Then, again using the technique of block multiplication, we have

$$G_2 G_1 A_1 = G_2 \begin{bmatrix} 1 & \alpha_1^{-1} b_1^T \\ 0 & A_2 \end{bmatrix} = \begin{bmatrix} 1 & \beta_2 & b_{11}^T \\ 0 & 1 & \alpha_2^{-1} b_2^T \\ 0 & 0 & B_2 - \alpha_2^{-1} a_2 b_2^T \end{bmatrix},$$

where

$$\alpha_1^{-1} b_1 = \begin{bmatrix} \beta_2 \\ b_{11}^T \end{bmatrix}.$$

Continuing in this way (assuming the *pivots* α_k are all non-zero), we obtain a matrix

$$U = G_n G_{(n-1)} \cdots G_1 A_1$$

that must be upper triangular with ones on the diagonal! This is simply the matrix U in the LU factorization of the matrix A_1. The associated matrix L is given by $L = (G_n G_{(n-1)} \cdots G_1)^{-1}$.

SAMPLE PROBLEMS

1. Consider the system

$$\begin{array}{rcrcrcl} 4x_1 & & & - & x_3 & = & 200 \\ 9x_1 & + & x_2 & - & x_3 & = & 200 \\ 7x_1 & - & x_2 & + & 2x_3 & = & 200\,. \end{array}$$

   (a) Write the augmented matrix corresponding to this system.

   (b) Reduce the augmented system in part (a) to echelon form.

   (c) Describe the set of solutions to the given system.

2. Represent the linear span of the four vectors

$$x_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -2 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 2 \\ 1 \\ 7 \\ 1 \end{bmatrix}, \quad \text{and} \quad x_4 = \begin{bmatrix} 3 \\ -2 \\ 0 \\ 5 \end{bmatrix},$$

   as the range space of some matrix.

3. Compute a basis for $\operatorname{nul}(A^T)^\perp$ where A is given by

$$A = \begin{bmatrix} 1 & -1 & 2 & 3 \\ 0 & 1 & 1 & -2 \\ 2 & 1 & 7 & 0 \\ 1 & -2 & 1 & 5 \end{bmatrix}.$$

4. Find the inverse of the matrix $B = \begin{pmatrix} 1 & 2 & 0 \\ -1 & -4 & 1 \\ 0 & 2 & 1 \end{pmatrix}$.

5. Solve the following system of linear equations

$$\begin{array}{rcrcrcl} x_1 & + & 2x_2 & & & = & 1 \\ -x_1 & - & 4x_2 & + & x_3 & = & 2 \\ & & 2x_2 & + & x_3 & = & 0. \end{array}$$

6. Determine whether the following system of linear equations has a solution or not.

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -2 \\ 0 \end{pmatrix}.$$

7. Find a 2 by 2 square matrix B satisfying

$$A = B \cdot C,$$

where $A = \begin{pmatrix} 1 & 3 & 0 \\ 2 & 1 & 1 \end{pmatrix}$ and $C = \begin{pmatrix} -1 & -3 & 0 \\ 8 & 9 & 3 \end{pmatrix}$.

8. What is the Gaussian elimination matrix for the vector

$$v = \begin{bmatrix} 2 \\ -10 \\ 16 \\ 2 \end{bmatrix}?$$

9. Show that the Gaussian elimination matrix for the vector

$$v = \begin{bmatrix} \alpha \\ a \end{bmatrix}$$

where $\alpha \in I\!R$ is non-zero and $b \in I\!R^{(n-1)}$ is non-singular by providing an expression for its inverse.

10. Show that the product of two lower triangular $n \times n$ matrices is always a lower triangular matrix.

11. Show that the inverse of a non-singular lower triangular matrix is always lower triangular.

12. A *Housholder transformation* on $I\!R^n$ is any $n \times n$ matrix of the form

$$P = I - 2\frac{vv^T}{v^T v}$$

for some non-zero vector $v \in I\!R^n$. Given any two vectors u and w in $I\!R^n$ such that $\|u\| = \|w\|$ and $u \neq w$, construct a Housholder transformation P such that $Pu = w$. For example, if

$$u = \begin{bmatrix} 3 \\ 1 \\ 5 \\ 1 \end{bmatrix} \quad \text{and} \quad w = \begin{bmatrix} -6 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

then

$$v = \begin{bmatrix} 9 \\ 1 \\ 5 \\ 1 \end{bmatrix}.$$