**Math 128a - Homework 9 - Due May 9**

1) Let $A$ be $m$-by-$n$, $B$ be $n$-by-$k$ and $C$ be $m$-by-$k$, and $x$ be $k$-by-1. Define matrix vector multiplication in the usual way: $y = B \cdot x$ is $n$-by-1 with $y_i = \sum_{j=1}^{k} B_{i,j} \cdot x_j$. Now suppose that $A \cdot (B \cdot x) = C \cdot x$ for all $k$-by-1 vectors $x$. Show that $C$ must be given by

$$C_{i,j} = \sum_{m=1}^{n} A_{i,m} \cdot B_{m,j}$$

This explains why matrix-matrix multiplication is defined the way it is.

2) Let $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ be $m$-by-$n$, where $A_{ij}$ is $m_i$-by-$n_j$ (we clearly assume that $m = m_1 + m_2$ and $n = n_1 + n_2$. $A$ is sometimes called a "block matrix". Similarly, let $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$ be $n$-by-$k$, where $B_{ij}$ is $n_i$-by-$k_j$ (we clearly assume that $k = k_1 + k_2$). Prove the following "block matrix multiplication formula"

$$A \cdot B = \begin{bmatrix} A_{11} \cdot B_{11} + A_{12} \cdot B_{21} & A_{11} \cdot B_{12} + A_{12} \cdot B_{22} \\ A_{21} \cdot B_{11} + A_{22} \cdot B_{21} & A_{21} \cdot B_{12} + A_{22} \cdot B_{22} \end{bmatrix}$$

This basically says that if the entries of $A$ and $B$ are themselves matrices with appropriate dimensions, the usual matrix multiplication formula works by substituting the blocks into the formula.

3) Assuming that $A$ is square and nonsingular, show that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix} \cdot \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix}$$

where all the blocks have dimensions as in the last problem. $D - CA^{-1}B$ is called the *Schur complement* of $A$, and shows up as an intermediate result in Gaussian elimination. Hint: use the last result of the last question.

4) Prove the *Sherman-Morrison formula*: If $A$ is nonsingular, $u$ and $v$ are column vectors such that $A + u \cdot v^T$ is nonsingular, then $(A + u \cdot v^T)^{-1} = A^{-1} - A^{-1} \cdot u \cdot v^T \cdot A^{-1}/(1 + v^T \cdot A^{-1} \cdot u)$ is an explicit expression for the inverse of $A + u \cdot v^T$. Hint: Multiply by $A + u \cdot v^T$ and simplify.

5) Suppose that we have done Gaussian elimination on $A$, so that solving $Ax = b$ for a new $b$ costs just $O(n^2)$, since we can reuse $A$'s $L$ and $U$ factors. Show that the following algorithm solves $(A + u \cdot v^T)x = b$ in $O(n^2)$ time. This is much cheaper than doing Gaussian elimination on $A + u \cdot v^T$, which would cost $2/3n^3$. This is useful because it is common to have to solve several different systems of linear equations where the matrices differ just by adding $u \cdot v^T$ for some columns vectors $u$ and $v$. Hint: Use the Sherman-Morrison formula.

1) Solve $Az = b$ for $z$
2) Solve $Ay = u$ for $y$
3) Compute $\alpha = v^T y$ (a dot product)
4) Compute $\beta = v^T z$ (a dot product)
5) Compute $x = z - \frac{\beta}{1+\alpha} y$ (adding multiple of one vector to another)

6) Suppose that we need to solve $Ax_i = b_i$ for $i = 1, ..., m$, i.e. for $m$ vectors $b_i$. There are two obvious algorithms for this:

Algorithm 1
Use Gaussian elimination to compute the L and U factors of $A$
For $i = 1$ to $m$
    Use the $L$ and $U$ factors to solve $Ax_i = b_i$
end

Algorithm 2
Use Gaussian elimination to compute the L and U factors of $A$
For $i = 1$ to $n$ ... compute the inverse of $A$
    Use the $L$ and $U$ factors to solve $Ay_i = e_i$, where $e_i$ is the $i$-th column of $I$
end
... Note that $A^{-1} = [y_1, y_2, ..., y_n]$
for $i = 1$ to $m$
    $x_i = A^{-1} \cdot b_i$ ... matrix-vector multiplication
end

Explain why the $n$-by-$n$ matrix $[y_1, ..., y_n]$ gotten by putting the vectors $y_i$ together in matrix is $A^{-1}$. Count the operations for each algorithm (your answers should look like $c_1 n^3 + c_2 mn^2 + O(n^2)$, where you need to figure out the constants $c_1$ and $c_2$). Count all operations (multiplication, addition, subtraction and division) equally. Conclude that Algorithm 1 is faster than Algorithm 2. This means that you should never compute an explicit inverse of $A$ to solve linear systems of equations, no matter how many of them you have; instead you should always use $A$'s $L$ and $U$ factors.