Mathematics 189-133B, Winter 2003
Vectors, Matrices and Geometry

A proof (in one place) of the Fundamental Theorem of Invertible Matrices, (version 3 on p. 290 of Poole, and then some). This proof is not entirely self-contained, but we try to make explicit what is used.

We will, in particular, be using without further mention the basic properties of matrix operations (associativity of multiplication and distributivity especially). We will also use the basic rules for row-reduction: a solution to $A\vec{x} = \vec{b}$ is the same as a solution to $A^*\vec{x} = \vec{b}^*$ whenever $(A|\vec{b})$ can be transformed to $(A^*|\vec{b}^*)$ by row operations (in case $\vec{b} = \vec{0}$, we must then have $\vec{b}^* = \vec{0}$, too); there is a unique reduced row-echelon form matrix R which can be obtained from A by the usual row operations and it is either the identity matrix, or has at least one row of zeroes (as it will be square); A^* (or $(A^*|\vec{b}^*)$) is obtained from A (or $(A|\vec{b})$) by a single row operation if and only if there is an elementary matrix E such that $EA = A^*$ (or $E(A|\vec{b}) = (A^*|\vec{b}^*)$).

In the following, A is an $n \times n$ matrix, and any vectors mentioned are vectors in $\mathcal{R}^n$. Of course, in all parts especially (8), (9) and (10) we are thinking of R^n as consisting of column vectors (as usual), but in (11), (12) and (13) we imagine the elements of R^n are row vectors.

The following are equivalent:

1. A is invertible;
2. $A\vec{x} = \vec{b}$ has a *unique* solution for each $\vec{b} \in R^n$;
   (a) $A\vec{x} = \vec{b}$ has *at least one* solution for each $\vec{b} \in R^n$;
   (b) $A\vec{x} = \vec{b}$ has *at most one* solution for each $\vec{b} \in R^n$;
3. $A\bar{x} = \bar{0}$ has only the trivial solution;
4. the reduced row-echelon form of A is I_n;
5. A is a product of elementary matrices;
6. $rank(A) = n$;
7. nullity(A)=0;
8. the column vectors of A are linearly independent;
9. the column vectors of A span R^n;
10. the column vectors of A form a basis for R^n;
11. the row vectors of A are linearly independent;
12. the row vectors of A span R^n;
13. the row vectors of A form a basis for R^n;

14. $det(A) \neq 0$;

15. 0 is not an eigenvalue of A;

16. A has a *right inverse* — that is, there is an $n \times n$ matrix B such that $AB = I_n$;

17. A has a *left inverse* — that is, there is an $n \times n$ matrix C such that $CA = I_n$;

18. A^T is invertible.

We start the proof with (1)⇒(2)⇒(2(b))⇒(3)⇒(4)⇒(5)⇒(1).

(1)⇒(2). Suppose that A is invertible. Consider any $\vec{x}_0$ such that $A\vec{x}_0 = \vec{b}$; multiplying on the left by A^{-1} gives that $A^{-1}(A\vec{x}_0) = A^{-1}\vec{b}$, so $\vec{x}_0 = I\vec{x}_0 = (A^{-1}A)\vec{x}_0 = A^{-1}\vec{b}$. This implies that the only possible solution is $A^{-1}\vec{b}$. But if $\vec{x}_0 = A^{-1}\vec{b}$, then $A\vec{x}_0 = A(A^{-1}\vec{b}) = (AA^{-1})\vec{b} = I\vec{b} = \vec{b}$, so that $\vec{x}_0 = A^{-1}\vec{b}$ *is* a solution.

(2)⇒(2(b)) is obvious.

(2(b))⇒(3) is clear, taking the special case $\vec{b} = \vec{0}$ of (2(b)) (and noting that the trivial solution $\vec{x} = \vec{0}$ is indeed a solution).

(3)⇒(4). Let R be the RREF form of A. So any solution to $R\vec{x} = \vec{0}$ is a solution to $A\vec{x} = \vec{0}$. If R is *not* I_n, it must have at least one row of zeroes and a column corresponding to a parameter; say it is the kth column. By choosing that parameter equal to 1 and any others equal to 0 (say) we can solve for a solution $\bar{x}$ with $x_k = 1$. In (3), we assume that the only solution to $A\vec{x} = \vec{0}$ is trivial, so the same is true of $R\vec{x} = \vec{0}$. Thus we have to have $R = I_n$, which is what (4) says.

(4)⇒(5). Row-reducing a matrix amounts to repeated multiplying it on the left by elementary matrices; in row-reducing A we get a sequence

$$A, E_1A, E_2(E_1A), E_3(E_2E_1A), \ldots, E_k(E_{k-1}\cdots E_1A) = R,$$

where R is the RREF of A. In case $R = I_n$, which (4) asserts, we use associativity repeatedly to get $(E_kE_{k-1}\cdots E_1)A = I$.

Now each E_j is not only invertible, but its inverse E_j^{-1} is also elementary. (Details later.) $E_1^{-1}E_2^{-1}\cdots E_k^{-1}$ is then a product of elementary matrices, and

$$A = IA = (E_1^{-1}E_2^{-1}\cdots E_k^{-1})((E_k\cdots E_2E_1)A) = (E_1^{-1}E_2^{-1}\cdots E_k^{-1})I = E_1^{-1}E_2^{-1}\cdots E_k^{-1}.$$

Those details: If E_j is like the identity matrix, except that there is an entry $c \neq 0$ in the (r, s)-place with $r \neq s$, then E_j^{-1} is almost identical, except that its (r, s)-entry is $-c$; if E_j is like the identity with two rows switched, then $E_j^{-1} = E_j$; finally, if E_j looks like the identity except one diagonal entry (say the (r, r)-entry) is $c \neq 0$ rather than 1 — in this case E_j^{-1} is like the identity except its (r, r)-entry is c^{-1}.

(5)⇒ (1). $A = E_1E_2 \cdots E_k$ is a product of elementary matrices, then A is a product of invertible matrices, and hence is itself invertible, with inverse $E_k^{-1} \cdots E_2^{-1}E_1^{-1}$. (Note the reversal of order.)

(2)⇒(2(a)) is clear.

(2(a))⇒(4). Let E_1 be any invertible matrix (maybe an elementary matrix). Then (2(a)) implies that we can always solve $(E_1A)\vec{x} = \vec{b}$ for any $\vec{b} \in R^n$. To see this, apply (2(b)) to find a solution $\vec{x}_0$ to $A\vec{x} = E_1^{-1}\vec{b}$; then

$$(E_1A)\vec{x}_0 = E_1(A\vec{x}_0) = E_1(E_1^{-1}\vec{b}) = (E_1E_1^{-1})\vec{b} = \vec{b}.$$

Repeating this, we see that $(E_k \cdots E_1)A)\vec{x} = \vec{b}$ always has a solution for any $\vec{b} \in R^n$ where the E_j's are elementary. If R is the RREF of A, $R = (E_k \cdots E_1)A$ for some elementary matrices $E_1,\ldots,E_k$. But the only matrix in RREF that satisfies condition (2(a)) is I_n since any other R has a row of zeroes and there's no solution to $R\vec{x} = \vec{e}_n$.

Now we have (1)⇔(2)⇔(2(a))⇔(2(b)) ⇔(3)⇔(4) ⇔(5).

Since the only RREF square matrix without a row of zeroes is the identity (4)⇔(6).

(3) and (7) are also restatements of each other. The nullity of A is just the dimension of the subspace $\{\vec{x} : A\vec{x} = \vec{0}\}$ of R^n. This dimension is 0 if and only if $\vec{0}$ is all alone in the subspace. So (3)⇔(7).

(3)⇒(8). Let $\vec{c}_1,\ldots,\vec{c}_n$ be the columns of A and suppose that $\vec{c}_1x_1 + \cdots + \vec{c}_nx_n = \vec{0}$. Then for $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ we have $A\vec{x} = \vec{0}$. Then (3) tells us that $\vec{x} = 0$. We conclude that$\{\vec{c}_1, \ldots, \vec{c}_n\}$ is independent, which is (8).

(8)⇒(3). Again, let $\vec{c}_1,\ldots,\vec{c}_n$ be the columns of A. If $A\vec{x} = \vec{0}$, where $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, then $\vec{c}_1x_1 + \cdots + \vec{c}_nx_n = \vec{0}$. By (8), we must have $x_1 = \cdots = x_n = 0$ and so $\vec{x} = \vec{0}$; this is (3).

(2(a))⇒(9). Let $\vec{b} \in R^n$ be arbitrary and let $\vec{c}_1,\ldots, \vec{c}_n$ be the columns of A. If $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ solves $A\vec{x} = \vec{b}$, which is possible by (2(a)), then $\vec{c}_1x_1+\cdots+\vec{c}_nx_n = \vec{b}$ and $\vec{b} \in span(\{\vec{c}_1, \ldots, \vec{c}_n\}$. Since $\vec{b}$ is arbitrary, $R^n = span(\{\vec{c}_1, \ldots, \vec{c}_n\}$.

(9)⇒(2(a)). Let $\vec{b} \in R^n$ be arbitrary and let $\vec{c}_1,\ldots, \vec{c}_n$ be the columns of A. Since $span(\{\vec{c}_1, \ldots, \vec{c}_n\}) = R^n$, there are $x_1,\ldots, x_n$ in R such that $\vec{c}_1x_1 + \cdots + \vec{c}_nx_n = \vec{b}$. Then $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ solves $A\vec{x} = \vec{b}$, verifying (2(a)).

Clearly, (10)⇒(8), and (10)⇒(9). By the above, (8)⇔(3)⇔(2(a)) ⇔(9). So if one of (8) and/or (9) is true, so is the other, and thus (8) by itself implies

(10), and (9) by itself also implies (10). At this stage, we have the equivalence of (1) through (10) proved (including (2(a)) and (2(b)).)

Since the rows of A are the columns of A^T, (11) is equivalent to (8), except for A^T instead of A. For the same reason, by what we have done, (11), (12), and (13) are each equivalent to (18), the statement that A^T is invertible. To show that (11), (12) and (13) are each equivalent to (1), then, it is sufficient to show that (1)$\Leftrightarrow$(18).

(1)$\Rightarrow$(18). Assuming (1), There is a matrix M such that $AM = MA = I$. So $I = I^T = (AM)^T = M^T A^T$ and $I = I^T = (MA)^T = A^T M^T$ and M^T is an inverse for A^T, which is (18).

(18)$\Rightarrow$(1). Suppose that $A^T Q = QA^T = I$. Then $I = I^T = (A^T Q)^T = Q^T (A^T)^T = Q^T A$ and $I = I^T = (QA^T)^T = (A^T)^T Q^T = AQ^T$, so that Q^T is an inverse for A and (1) is then true.

(It is instructive to prove the equivalence of (11), (12) and (13) to the earlier conditions on A without going through the transpose. One way is to note that row-reducing A does not change whether or not A is invertible — if E is elementary, then EA is invertible if and only if A is — and also row-reducing doesn't change the truth or falsity of statements (11), (12) and/or (13). Then use the fact that the only RREF matrix that is invertible is I_n, but it is also the sole square matricein RREF whose rows are independent, and the only one whose rows span R^n regarded as a collection of row vectors. We leave the details.)

To prove the equivalence of (4) and (14), we cheat by using some other properties of determinants. Specifically, we use

1. Row-reducing a matrix does not change whether or not its determinant is zero. [In more, detail, replacing row $\vec{r}_j$ by $\vec{r}_j + a\vec{r}_k$, where $\vec{r}_k$ is some other row, doesn't alter the determinant at all; switching two rows changes the sign fo the determinant; and multiplying $\vec{r}_j$ by the nonzero constant c has the effect of multiplying the determinant by c.]

2. The only $n \times n$ matrix in RREF with determinant different from 0 is the identity matrix.

So $det(A) \neq 0 \Leftrightarrow$ the RREF of A has determinant $\neq 0 \Leftrightarrow$ the RREF of A is I. That is, (4)$\Leftrightarrow$(14).

To finish the proof we need to fit (15), (16) and (17) into the picture. We do this by proving $\neg$(3) $\Rightarrow$ $\neg$(15), $\neg$(15) $\Rightarrow$ $\neg$(3), (16)$\Rightarrow$(2(a)), and then (17)$\Rightarrow$(3). This is enough as (1)$\Rightarrow$(16) and (1)$\Rightarrow$(17) are clear.

$\neg$(3) $\Rightarrow$ $\neg$(15). Suppose that 0 is an eigenvalue of A. Then there is an eigenvalue $\vec{v}$ for 0. Necessarily $\vec{v} \neq \vec{0}$, and $A\vec{v} = 0\vec{v} = \vec{0}$, so that there is a nontrivial solution to $A\vec{x} = \vec{0}$.

$\neg$(15) $\Rightarrow$ $\neg$(3). If $A\vec{v} = \vec{0}$ with $\vec{v} \neq \vec{0}$, then $A\vec{v} = 0\vec{v}$ and $\vec{v}$ is an eigenvector for the eigenvlaue 0.

(16)$\Rightarrow$(2(a)). Suppose that $AB = I$ and $\vec{b} \in R^n$; letting $\vec{x} = B\vec{b}$, we have $A\vec{x} = A(B\vec{b}) = (AB)\vec{b} = I\vec{b} = \vec{b}$, so for any $\vec{b}$ we have a solution to $A\vec{x} = \vec{b}$.

(17)$\Rightarrow$(3). Suppose that $CA = I$. If $A\vec{x} = \vec{0}$, then $C(A\vec{x}) = C(\vec{0}) = \vec{0}$, so $\vec{x} = I\vec{x} = (CA)\vec{x} = \vec{0}$; thus (3) holds.