14.102, Math for Economists
Fall 2005
Lecture Notes, 9/20/2005

These notes are primarily based on those written by George Marios Angeletos for the Harvard Math Camp in 1999 and 2000, and updated by Stavros Panageas for the MIT Math for Economists Course in 2002. I have made only minor changes to the order of presentation, and added some material from Guido Kuersteiner's notes on linear algebra for 14.381 in 2002. The usual disclaimer applies; questions and comments are welcome.

Nathan Barczi
nab@mit.edu

4 Span, Basis, and Rank

4.1 Linear Combinations

Fix m and n; take a set of n vectors $\{a_j\} = \{a_1, a_2, ..., a_n\}$, where a_j is a (column) vector in $\mathbb{R}^m$; take n real numbers $\{x_j\} = \{x_1, ..., x_n\}$; and construct a vector y in $\mathbb{R}^m$ as the sum of the a_j's weighted by the corresponding x_j's:

$$y = x_1 a_1 + x_2 a_2 + ... + x_n a_n$$

Then the so-constructed vector y is called a **linear combination** of the a_j's.

If we let $A = [a_j]$ be the $m \times n$ matrix with columns the vectors a_j's and x the n-dimensional vector $[x_j]$, then we can write y as

$$y = Ax = \sum_{j=1}^{n} x_j a_j$$

Thus, Ax is a linear combination of the columns of A.

Notice that the dimension of the vector $y = Ax$ is the same as of that of any column a_j. That is, y belongs to the same vector space as the a_j's.

4.2 Linear Dependence/Independence

Consider a set of n vectors in $\mathbb{R}^m$, $\{a_j\} = \{a_1, a_2, ..., a_n\}$. These vectors are called **linearly dependent** if any one of them can be written as a linear combination of the rest. They are otherwise called **linearly independent**.

Definition 102 *Letting $A = [a_j]$, the vectors a_j's are **linearly independent** if*

$$Ax = 0 \Rightarrow x = 0$$

*They are **linearly dependent** if*

$$\exists\, x \in \mathbb{R}^m \; s.t.\, x \neq 0 \; and \; Ax = 0$$

Recall that two vectors x and z are called **orthogonal** iff $z'x = 0$. Thus, $Ax = 0$ means that x is orthogonal to all the rows of A. Similarly, $A'x = 0$ (for x now in $\mathbb{R}^m$ since A' is $n \times m$) means that x is orthogonal to all the columns a_j's of A.

Note that the maximum number of linearly independent vectors that we can have in any given collection of n-dimensional vectors is n. To see this, note that the equation $Ax = 0$ can be thought of as a system of n equations (where n is the dimension of A), with the number of unknowns being equal to the dimension of x, which is also the number of columns (vectors) in A. But it is a fact that a homogenous system of equations (i.e., one with zeros on the right-hand side of every equation) with more unknowns than equations must have infinitely many solutions, all but one of which are nonzero. On the other hand, we know that we can write n linearly independent vectors of dimension n - the n-dimensional identity matrix consists of just such a collection.

4.3 The Span and the Nullspace of a Matrix, and Linear Projections

Consider an $m \times n$ matrix $A = [a_j]$, with a_j denoting its typical column. Consider then the set of all possible linear combinations of the a_j's. This set is called the span of the a_j's, or the **column span** of A.

Definition 103 *The **(column) span** of an $m \times n$ matrix A is*

$$\begin{aligned} S(A) &\equiv S[a_1, ..., a_n] \equiv \\ &\equiv \left\{ y \in \mathbb{R}^m \mid y = Ax = \sum_{j=1}^n x_j a_j \ \ for\, some\, x = [x_j] \in \mathbb{R}^n \right\} \end{aligned}$$

On the other hand, we define the nullspace of A as the set of all vectors that are orthogonal to the rows of A.

Definition 104 *The **nullspace** or **Kernel** of an $m \times n$ matrix A is*

$$\begin{aligned} N(A) &\equiv N[a_1, ..., a_n] \equiv \\ &\equiv \{ x \in \mathbb{R}^n \mid Ax = 0 \} \end{aligned}$$

Exercise 105 *Given A, show that $S(B) \subseteq S(A)$ and $N(A') \subseteq N(B')$ whenever $B = AX$ for some matrix X. What is the geometric interpretation?*

Notice that $N(A')$ is the set of all vectors that are orthogonal to the a_j's. Thus,

$$z \in N(A') \Leftrightarrow z \perp S(A)$$

which means that $N(A')$ is the **orthogonal complement** subspace of $S(A)$. That is,[5]

$$S(A) + N(A') = \mathbb{R}^m$$

Indeed:

[5] Recall that, for any sets X, Y, their sum is defined as $X + Y \equiv \{ z \mid z = x + y, \text{ some } x \in X, y \in Y \}$.

Exercise 106 *Given an $m \times n$ matrix A, show that $S(A), N(A)$ and $N(A')$ are all linear subspaces. Show further that $S(A)$ and $N(A')$ are orthogonal subspaces, in the sense that $z \in S(A), u \in N(A') \Rightarrow z'u = 0$. Show further that $S(A) + N(A') = \mathbb{R}^m$, in the sense that for every $y \in \mathbb{R}^m$ there are vectors $z \in S(A)$ and $u \in N(A')$ such that $y = z + u$.*

Remark: z is then called the **(linear) projection** of y **on** $S(A)$, or the regression of y on A, and u is called the **residual**, or the projection **off** $S(A)$. Does this reminds you something relevant to econometrics?

The last results are thus summarized in the following:

Lemma 107 *$S(A)$ and $N(A')$ form an orthogonal partition for $\mathbb{R}^m$; that is,*

$$S(A) \perp N(A') \quad \text{and} \quad S(A) + N(A') = \mathbb{R}^m$$

4.4 Basis of a Vector Space

Let $\mathbb{X}$ be a vector space. We say that the set of vectors $\{a_1, ..., a_n\} \subset \mathbb{X}$, or the matrix $A = [a_j]$, **spans** $\mathbb{X}$ iff $S(a_1, ..., a_n) = S(A) = \mathbb{X}$.

If A spans $\mathbb{X}$, it must be the case that any $x \in \mathbb{X}$ can be written as a linear combination of the a_j's. That is, for any $x \in \mathbb{R}^n$, there are real numbers $\{c_1, ..., c_n\} \subset \mathbb{R}$, or $c \in \mathbb{R}^n$, such that

$$x = c_1 a_1 + ... + c_n a_n, \quad \text{or} \quad x = Ac$$

There may be only one or many c such that $x = Ac$. But if for each $x \in \mathbb{X}$ there is only a unique c such that $x = Ac$, then c_j is called the j-th coordinate of x with respect to A. And then A is indeed a basis for $\mathbb{X}$.

Definition 108 *A* ***basis*** *for a vector space $\mathbb{X}$ is any set of linearly independent vectors, or a matrix, that spans the whole $\mathbb{X}$.*

Example 109 *In $\mathbb{R}^n$, the usual basis is given by $\{e_1, ..., e_n\}$ where e_i is a vector with a unit in the i-th position and zeros elsewhere; alternatively, the $n \times n$ identity matrix*

$$I = \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & ... & 1 \end{bmatrix}$$

is a basis for $\mathbb{R}^n$. If $x = (x_1, ..., x_n) \in \mathbb{R}^n$, then x_j are simply the coordinates of x with respect to I; that is,

$$x = x_1 e_1 + x_2 e_2 + ... + x_n e_n = Ix$$

This means that $\{e_1, ..., e_n\}$, or I, spans $\mathbb{R}^n$. And trivially, the e_j's are linearly independent, because

$$xI = x, \quad or \ x_1 e_1 + ... + x_n e_n = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

and thus $xI = 0 \Rightarrow x = 0$.

Observe that any space may admit many-many different bases! For instance:

Exercise 110 *Show that if $\{e_j\}$ is a basis for $\mathbb{X}$, then so is $\{f_j\} = \{\mu e_j\}$ for any scalar $\mu \neq 0$.*

And a bit less trivial:

Exercise 111 *Suppose $\{e_j\}$ is a basis for $\mathbb{X}$; let $P = [p_{ij}]$ be any nonsingular $n \times n$ matrix, and let $f_j = \sum_i p_{ij} e_i$. Show then that $\{f_j\}$ is a basis for $\mathbb{X}$ as well.*

In other words,

Lemma 112 *If E is a basis for $\mathbb{X}$, then so is $F = EP$ for any nonsingular P.*

Notice that, with F so-constructed, P is then simply the coordinates of the basis F with respect to the basis E.

Example 113 *For instance, all of the following matrices are bases for $\mathbb{R}^2$:*

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$$

provided $\alpha\delta - \beta\gamma \neq 0$.

Exercise 114 *Can you figure out what are the coordinates of $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ with respect to the above alternative bases? What about any $x = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]$?*

Exercise 115 *Characterize the set of all bases for the real line, $\mathbb{R}$. Do the same for $\mathbb{R}^n$. Persuade yourself that this is the set of all nonsingular $n \times n$ matrices.*

In the above example for $\mathbb{R}^2$, we found many different bases, but they all had something in common: They were all made of just 2 vectors, and we know well that 2 is the dimension of $\mathbb{R}^2$. But, what is the dimension of a vector space, and is this unique despite the multiplicity of bases? In answering this the following helps:

Lemma 116 *Let $\{e_j\} = \{e_1, ..., e_n\}$ be a basis for $\mathbb{X}$, and let $\{b_j\} = \{b_1, ..., b_m\}$ be any set of vectors with $m > n$. Then $\{b_j\}$ can not be linearly independent.*

Exercise 117 *Can you prove this? Here is a hint: Write $b_j = \sum_i c_{ij} e_i$ for some $\{c_{ij}\}$. Let $C = [c_{ij}]$ and let $x \neq 0$ be some solution to $Cx = 0$. [Which lemma/proposition ensures that such a solution exists?] Use that to show*

$$\sum_i \lambda_i e_i = 0 \ \ \textit{for} \ \ \lambda_i = \left(\sum_j x_j c_{ij} \right) \neq 0 \ \forall i$$

But isn't that a contradiction, which indeed completes the proof?

From the last lemma it follows that all bases of a given space will have the same number of elements. Therefore, we can unambiguously define:

Definition 118 *The* ***dimension*** *of a vector space* $\mathbb{X}$, $\dim(\mathbb{X})$, *is the number of elements in any of its bases. On the other hand, if such a basis with finite elements does not exist, then the space is infinite dimensional.*

Exercise 119 *Consider the space of all sequences of real numbers. Is that a vector space? Can you think of a basis for it? Let* e_j *be a sequence that has unit in its j'th position and zero in all other positions — e.g.,* $e_2 = \{0, 1, 0, 0, ...\}$ *— and consider the set* $\{e_j | j = 1, 2, ...\}$. *Is that a basis for the space of sequences? What is its dimension?*

4.5 The Rank and the Nullity of a Matrix

The rank of matrix $A = [a_j]$ is defined as the maximum number of independent columns a_j of this matrix. In particular,

Definition 120 *The* ***rank*** *of a matrix A is the dimension of its span. The* ***nullity*** *of A is the dimension of its nullspace. That is,*

$$rank(A) \equiv \dim(S(A)) \quad and \quad null(A) \equiv \dim(N(A))$$

A useful result to keep in mind is the following:

Lemma 121 *Let any matrix A, and* A' *its transpose. Then, the rank of A and* A' *coincide:*

$$rank(A) = rank(A')$$

This simply means that a matrix always has as many linearly independent columns as linearly independent rows. Equivalently, a matrix and its transpose span subspaces of the same dimension.

But, is there any relation between the rank and the nullity of a matrix? There is indeed, and this consistitutes the '**fundamental theorem of linear algebra**':

Theorem 122 *Let any* $m \times n$ *matrix* $A = [a_j]$, *with n columns* $a_j \in \mathbb{R}^m$. *Then, its rank and its nullity sum up to n:*

$$rank(A) + null(A) \; = \; n \; = \; \#\{a_j\}$$

Exercise 123 *Here is a sketch of the proof [not an easy one]; you have to fill the details: Let* $k = null(A) \equiv \dim(N(A))$. *[Check that* $k \leq m$.*] Take a basis* $\{e_1, ..., e_k, e_{k+1}, ..., e_n\}$ *of* $\mathbb{R}^n$ *such that* $\{e_1, ..., e_k\}$ *is a basis for* $N(A)$. *For any* $x \in \mathbb{R}^n$, *the are* $\lambda_1, ..., \lambda_n \in \mathbb{R}$ *such that*

$$x = \lambda_1 e_1 + ... + \lambda_n e_n$$

and

$$Ax = \lambda_{k+1} Ae_{k+1} + ... + \lambda_n Ae_n$$

because

$$\{e_1, ..., e_k\} \subset N(A) \Rightarrow Ae_1 = ... = Ae_k = 0$$

Show further that the set $\{Ae_{k+1}, ..., Ae_n\}$ *is linearly independent as well: Assume not and get a contradiction that* $\{e_1, ..., e_n\}$ *would then be linearly dependent. Conclude that* $\{Ae_{k+1}, ..., Ae_n\}$ *forms a basis for* $S(A)$*. Notice that* $\{Ae_{k+1}, ..., Ae_n\}$ *has* $n-k$ *elements, and thus* $rank(A) = \dim(S(A)) = n-k = n - null(A)$*. QED*

A related result is the following:

Exercise 124 *Using the last theorem and the previous lemma, show that*

$$\begin{aligned} rank(A) + null(A') &= m \\ null(A) - null(A') &= n - m \end{aligned}$$

Remark: Recall that $S(A)$ and $N(A')$ form a partition (an orthogonal one, indeed) of $\mathbb{R}^m$. It is not thus surprising that $\dim(S(A)) + \dim(N(A')) = \dim(\mathbb{R}^m)$, or $rank(A)+null(A') = m$. From a 'transpose' view, $S(A')$ and $N(A)$ form a partition of $\mathbb{R}^n$, and thus $rank(A')+null(A) = n$, or $rank(A)+null(A) = n$, using the fact that $rank(A') = rank(A)$. Does this provide you with a clear geometric intuition for the above theorem?

4.6 Nonsingularity and Matrix Inversion

Definition 125 *A square matrix* A *of dimension* $n \times n$ *is* ***nonsingular*** *if* $rank(A) = n$*. Equivalently,* A *is nonsingular if* $null(A) = 0$*.*

Lemma 126 *If* A *is nonsingular then there exists a nonsingular matrix* A^{-1} *such that*

$$AA^{-1} = A^{-1}A = I$$

where I *is the identity matrix.*

Proof. *Recall that if* A *is nonsingular, then its columns are linearly independent. Therefore, the equation* $Ax = c$ *has a unique solution* x*. In particular,* $Ax = e_i$ *has a unique solution, where* e_i *is the* $n \times 1$ *column vector with a* 1 *as its* i^{th} *element and zeroes elsewhere. We can stack* n *such equations* $(Ax_1 = e_1, Ax_2 = e_2, ..., Ax_n = e_n)$ *to show that* $AX = I$ *is satisfied by a unique matrix* X*; this matrix is a right inverse of* A*. To show that* A *has a left inverse, note that* A *nonsingular implies that the rows of* A *are linearly independent as well, and so the equation* $yA = e_i'$*, where* e_i' *is the transpose of* e_i*, has a unique solution* x*, where* y *is a row vector. Continuing as before we see that* $YA = I$ *is satisfied by a unique matrix* Y*, and this is a left inverse of* A*.*

Finally, to show that $X = Y$ *(and that both can therefore truly be called* A^{-1}*), we will show that if* $AX = YA = I$*, then* $X = Y$*. To see this, suppose it's not* *true; $AX = YA = I$ but $X \neq Y$. Then premultiply the left-hand side of the inequality by YA and post-multiply the right-hand side by AX (which changes nothing, since both are equal to the identity matrix) and we have $YAX \neq YAX$, which is false and establishes the assertion.* ■

The following facts about matrix inverses are useful (assuming invertibility):

Lemma 127 $(AB)^{-1} = B^{-1}A^{-1}$
Proof. *Observe that if we postmultiply AB by its inverse $(AB)^{-1}$ we get I by the definition of matrix inverse, and similarly if we premultiply AB by $B^{-1}A^{-1}$ we get I, so we can apply the previous lemma to conclude that $(AB)^{-1} = B^{-1}A^{-1}$.*

Lemma 128 $(A')^{-1} = (A^{-1})'$
Proof. *We want to show that $A'(A^{-1})' = (A^{-1})'A' = I$. But this follows directly from our transpose rules, taking the transpose on both sides of the equalities $AA^{-1} = I$ (which gives $(A^{-1})'A' = I$) and $A^{-1}A = I$ (which gives $A'(A^{-1})' = I$).* ■

■

Note that the fact shown in the first of the two proofs above implies that the left-hand inverse and the right-hand inverse of a matrix are the same, so that we can talk meaningfully of *the* inverse of a nonsingular square matrix

4.7 Powers of a Matrix

For any square matrix A we may define its **k-th power**, denoted by A^k, for any $k \in \mathbb{N}$. A^k is defined inductively by $A^0 \equiv I$ and $A^k \equiv AA^{k-1}$ for $k \in \mathbb{N}_* = \{1, 2, ...\}$.

If A is invertible, then we also define $A^{-k} \equiv (A^{-1})^k$, for any $k \in \mathbb{N}$. Actually it can be shown that $A^{-k} = (A^k)^{-1}$.(See the exercise below)

Remark: So far we have not defined A^k for a non-integer but real number k. This will become possible for nonsingular and symmetric, or any diagonalizable matrices, but we have to defer till the point we discuss matrix eigensystem and diagonalization..

Exercise 129 *Use the properties of transpose and inverse to prove that*
1.)$A^{-k} = (A^k)^{-1}$
2.)Consider the matrix: $Z = X(X'X)^{-1}X'$ where X an arbitrary $m \times n$ matrix. Show that Z is symmetric. Also show that $ZZ = Z$.

4.8 The Determinant of a Matrix

The **determinant**, $\det(A)$ or $|A|$, of a matrix A is defined iff A is square.

Definition 130 *The determinant of a square matrix A with dimension $n \times n$ is a mapping $A \to |A|$ such that*

i) $|\cdot|$ *is linear in each row of* A.
ii) if $rank(A) < n$ *then* $|A| = 0$ *and vice versa*
iii) $|I| = 1$

We can view the matrix A as a collection of n row vectors $\{a_1, a_2, ..., a_n\}$ where each $a_i \in \mathbb{R}^n$. The determinant is then a function mapping the set of vectors $\{a_1, a_2, ..., a_n\}$ into $\mathbb{R}$. (We will not prove this here, but this mapping exists and it is unique, for any square matrix A). We can write $D(a_1, a_2, ..., a_n) \to \mathbb{R}$. Property i) of the previous definition means that

$$D(a_1, a_2, ..., \lambda a_i + \mu a_i', ..., a_n) = \lambda D(a_1, a_2, ..., a_i, ..., a_n) + \mu D(a_1, a_2, ..., a_i', ..., a_n)$$

for any scalars $\lambda, \mu \in \mathbb{R}$ and any vectors $a_i, a_i' \in \mathbb{R}$. In particular note that $D(a_1, a_2, ..., \lambda a_i, ..., a_n) = \lambda D(a_1, a_2, ..., a_i, ..., a_n)$.

Alternatively, we can give an inductive definition, providing a computational algorithm for the determinant:

Definition 131 *Let an* $n \times n$ *matrix* $A = [a_{ij}]$. *If* $n = 1$, *and hence* $A = [a_{11}]$, *then* $|A| \equiv a_{11}$. *For any* $n > 1$, *we let*

$$|A| \equiv \sum_{j=1}^{n} (-1)^{j+1} a_{1j} |A_{1j}| = a_{11}|A_{11}| - a_{12}|A_{12}| + ... \pm a_{1n}|A_{1n}|$$

where $[a_{1j}]$ *is the first row of* A *and* A_{1j} *is the* $(n-1) \times (n-1)$ *submatrix constructed by erasing the first row and the* j*-th column of* A.

This definition implies for a 2×2 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

that $|A| = a_{11}a_{22} - a_{12}a_{21}$.

For a 3×3 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

the determinant can be computed by the **Sarrus Rule**, which works as follows: First border the matrix at its right with its first two columns;

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix}$$

then take the sum of the products of the elements along the parallels of the principal diagonal minus the sum of the products of the elements along the parallels to the other diagonal; this gives the determinant of A:

$$\begin{aligned} |A| &= (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) \\ &\quad -(a_{31}a_{22}a_{13} + a_{32}a_{23}a_{11} + a_{33}a_{21}a_{12}) \end{aligned}$$

Sometime useful in speeding up computation, the following lemma allows us to work out the determinant along any raw or column:

Lemma 132 *For any fixed $i = 1, ..., n$,*

$$|A| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} |A_{ij}| = (-1)^{i+1} \left(a_{i1}|A_{i1}| - a_{i2}|A_{i2}| + ... \pm a_{in}|A_{in}| \right)$$

and for any fixed $j = 1, ..., n$

$$|A| = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} |A_{ij}| = (-1)^{j+1} \left(a_{1j}|A_{1j}| - a_{2j}|A_{2j}| + ... \pm a_{nj}|A_{nj}| \right)$$

where $[a_{ij}]_{j=1,..,n}$ is the i-th raw of A, $[a_{ij}]_{i=1,..,n}$ is the j-th column of A, and A_{ij} is the $(n-1) \times (n-1)$ submatrix constructed by erasing the i-th row and the j-th column of A.

Finally, keep in mind the following:

Lemma 133 *1.)* $|AB| = |A||B|$
2.) $|A| = |A'|$
3.) $\left|A^{-1}\right| = 1/|A|$

The determinant of a matrix allows an interesting interpretation in terms of the surface (more generally volume) of the vectors comprising a matrix.

4.9 Matrix Inversion, part II

The ordinary test for invertibility of a matrix is whether its determinant vanishes or not:

Theorem 134 *A matrix A is* ***invertible*** *iff $|A| \neq 0$*

Provided $|A| \neq 0$, A^{-1} exists for sure. However, the computation of A^{-1} is generally a pain, especially when $n > 2$. An algorithm to compute the inverse of a given $n \times n$ matrix $A = [a_{ij}]$ works as follows:

- For $n = 1$, and hence $A = [a_{11}]$, then $A^{-1} = \left[\frac{1}{a_{11}}\right] = \left[\frac{1}{|A|}\right]$.
- For $n > 1$: Let A_{ij} denote the $(n-1) \times (n-1)$ submatrix constructed by erasing the i-th row and the j-th column of A. Let $|A_{ij}|$ be called the (i, j) **first-order minor** of A and define the (i, j) **cofactor** of A as

$$c_{ij} = (-1)^{i+j} |A_{ij}|$$

Notice by the way that

$$\begin{aligned} |A| &= \sum_i a_{ij} c_{ij} = \sum_j a_{ij} c_{ij} = \\ &= \sum_j (-1)^{i+j} a_{ij} |A_{ij}| \end{aligned}$$

For example,

$$|A| = a_{11}|A_{11}| - a_{12}|A_{12}| + ... + (-1)^{n+1}a_{1n}|A_{1n}|$$

Construct the $n \times n$ matrix $C = [c_{ij}]$ of all first-order cofactors of A and define the **adjoint** of A as the transpose of C,

$$adj A \equiv C' = [c_{ji}]$$

Then, the inverse of A is

$$\begin{aligned} A^{-1} &= \frac{1}{|A|} adj A = \\ &= \frac{1}{|A|} \begin{bmatrix} +|A_{11}| & -|A_{21}| & ... & (-1)^{n+1}|A_{n1}| \\ -|A_{12}| & +|A_{22}| & ... & (-1)^{n+2}|A_{n2}| \\ ... & ... & & ... \\ (-1)^{n+1}|A_{1n}| & (-1)^{n+2}|A_{1n}| & ... & +|A_{nn}| \end{bmatrix} \end{aligned}$$

This algorithm works pretty well for manual computation if $n = 2$ or 3.

Exercise 135 *Show that for a* 2×2 *matrix*

$$A = \begin{bmatrix} a & \beta \\ \gamma & \delta \end{bmatrix}$$

provided $|A| = \alpha\delta - \beta\gamma \neq 0$, *the inverse is*

$$A^{-1} = \frac{1}{\alpha\delta - \beta\gamma} \begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix}$$

But otherwise, thanks to modern technology, computers can be our 'slaves' in computing the inverse of a huge matrix. [6]

Finally,

Exercise 136 *Show that if A is real and invertible, then* A^{-1} *is real as well. What if A is complex but generally non-real? Define the inverse of A appropriately.*

4.10 Matrix Inversion and Linear Independence: So far, so good...

This is probably the most important characterization result to remember. ♠ A kind of six-in-one shampoo....

[6] Check out Mathematica, Matlab, Fortran, etc. ♠ *Up to the 70's a research assistanship would quite possibly involve manual computation of matrix inverses. In contemporary academic research, computers have substituted graduate students in this job. Nonetheless, graduate-student slavery is still prevailing, in other old or new forms!*

Theorem 137 *Let A be a $n \times n$ square matrix. Then, the following conditions are equivalent*
(i) A is nonsingular; i.e., A^{-1} exists; $\Leftrightarrow$
(ii) A has a non-zero determinant, $|A| \neq 0$; $\Leftrightarrow$
(iii) the columns $\{a_j\}$ of A are linearly independent; i.e., $Ax = 0 \Rightarrow x = 0$; $\Leftrightarrow$
(iv) A forms a basis for $\mathbb{R}^n$; $\Leftrightarrow$
(v) A spans the whole n-dimensional space, $S(A) = \mathbb{R}^n$; $\Leftrightarrow$
(vi) the kernel of A is null, $N(A) = \{0\}$.

Further, the following lemma will prove useful when we analyze linear equations systems:

Lemma 138 *Let $1 \leq m < n$ and an $m \times n$ matrix $A = [a_{ij}]$. Then there is $x \in \mathbb{R}^n, x \neq 0$ such that $Ax = 0$. This in turn implies $null(A) \equiv \dim(N(A)) \geq 1$.*

Remark: If we interpret $N(A)$ as the set of solutions to the system $Ax = 0$ [see next section for details], the above lemma says that, whenever there are more unknowns than equations ($n > m$), then the system $Ax = 0$ is underdetermined and admits a continuum of solutions.

Exercise 139 *Construct an $n \times n$ matrix $B = [b_{ij}]$ by $b_{ij} = a_{ij}$ for $i = 1, ..., m$ and $b_{ij} = 0$ for $i = m+1, ..., n$, $\forall j$; more compactly, $B = \begin{bmatrix} A \\ 0 \end{bmatrix}$. What is $|B|$? What does this imply for Bx? Notice that $Bx = \begin{bmatrix} Ax \\ 0 \end{bmatrix}$. Does this help you prove the above lemma?*