5. Basic Linear Lore. Much of linear algebra over a field K can be deduced from the following basic lemma in which, for brevity, a matrix will be called *strongly regular* if it is a product of addition or permutation type elementary matrices. In particular, such a matrix is square and has an explicit left and right inverse. On the other hand, a matrix A will be called *singular* if it has a non-zero kernel $\mathcal{N}(A)$. Obviously these two properties exclude one another.

LEMMA: Let A be an $m \times n$ matrix over K. Then there exist strongly regular matrices M and N such that

$$MAN = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \qquad \text{where} \qquad D = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_r \end{bmatrix} \qquad \text{with} \qquad d_i \neq 0.$$

Proof: Let $\alpha(M,N)$ stand for the entry in the first row and first column of MAN. If $\alpha(M,N) = 0$ for all M, N, then obviously $A = 0$, and we are finished. Otherwise there is a pair M_1, N_1 such that

$$M_1AN_1 = \begin{bmatrix} d_1 & X \\ Y & A' \end{bmatrix},$$

where A' is an $(m-1) \times (n-1)$-matrix, and $d_1 \neq 0$. Multiplying on the left by addition-type elementary matrices, we make $Y = 0$. Similarly, operating from the right, we modify N_1 to get $X = 0$. The proof is finished by induction.

THEOREM: Let A be an $m \times n$ matrix with $m \leq n$. Then $\mathcal{N}(A) = \{0\} \iff A$ is square and invertible.

Proof: For invertible M, N it is easy to see that A is non-singular if and only if MAN is. If the latter is as above, non-singularity clearly means $r = m = n$. But then $MAN = D$ is an invertible diagonal matrix, and $A = M^{-1}DN^{-1}$ is invertible.

COROLLARY: An independent subset of the span of r vectors cannot have more than r elements.

Proof: Suppose $W_1, \ldots, W_s$ are in the span of $V_1, \ldots, V_r$; say $W_j = a_{1j}V_1 + \cdots + a_{rj}V_r$, for $j = 1, \ldots, s$. Consider the linear combination

$$x_1W_1 + \cdots + x_sW_s = (a_{11}x_1 + \cdots + a_{1s}x_s)V_1 + \cdots + (a_{r1}x_1 + \cdots + a_{rs}x_s)V_s.$$

If $s > r$, our theorem guarantees the existence of a non-trivial s-tuple $x_1, \ldots, x_s$ such that all this is zero, because the matrix (a_{ij}) involved here has more columns than rows, hence must be singular.

Note: Let $\mathcal{V}$ be a subspace of K^n. By the Corollary, any two bases of $\mathcal{V}$ have the same cardinality $\dim \mathcal{V}$. Moreover, any independent $\{W_1, \cdots, W_s\} \subset \mathcal{V}$ is contained in a basis of $\mathcal{V}$.

To see this, start with $W_1, \ldots, W_s$ and keep adjoining more vectors $W_{s+1}, W_{s+2}, \ldots \in \mathcal{V}$ (if you can), while maintaining the independence of your collection. By the Corollary, this process cannot go beyond a total of n vectors. At some point, therefore, your set $\{W_1, \ldots, W_{s+p}\}$ must stop being enlargeable; i.e. any additional vector $V \in \mathcal{V}$ must be a linear combination of the ones you already have.

This result also shows that $\dim \mathcal{V}$ is a meaningful measure of the "size" of $\mathcal{V}$. More precisely, if $\mathcal{V}$ contains a smaller subspace $\mathcal{V}'$, we can enlarge a basis of $\mathcal{V}'$ to one of $\mathcal{V}$, thus proving that $\dim \mathcal{V}' < \dim \mathcal{V}$.

Exercise: To test your understanding of dimension, try to prove the following identities:

$$n - \dim \mathcal{N}(A) = \dim \mathcal{C}(A) = \dim \mathcal{R}(A),$$

where $\mathcal{C}$ and $\mathcal{R}$ denote the spans of the columns and of the rows, respectively. For the first, take a basis $\{W_1, \ldots, W_k\}$ of $\mathcal{N}(A)$ and extend it to one $\{W_1, \ldots, W_n\}$ of K^n; then show that $\{AW_{k+1}, \ldots, AW_n\}$ is a basis of $\mathcal{C}(A)$. For the second, note that both dimensions are invariant under elementary row and column operations, hence equal to those of $\mathcal{C}(MAN)$ and $\mathcal{R}(MAN)$.

6. Real Matrices. One of the most central theorems about real matrices is also one of the easiest to prove. Its geometric version says that *any real linear transformation has the effect of mapping some orthonormal basis of the domain onto an orthogonal subset of the range.* Here is a simple proof of the matrix version.

THEOREM: Let A be an $m \times n$ real matrix. Then there exist orthogonal matrices M and N such that

$$MAN = \begin{bmatrix} D & 0 \\ \\ 0 & 0 \end{bmatrix} \qquad \text{where} \qquad D = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_r \end{bmatrix} \qquad \text{with} \qquad d_i \geq d_{i+1} > 0.$$

Proof: Let $O(n)$ be the set of all $n \times n$ orthogonal matrices. For $M \in O(m)$ and $N \in O(n)$, let $\alpha(M, N)$ stand for the entry in the first row and first column of MAN. Let d_1 the greatest possible value occurring among these.

Since $O(m) \times O(n)$ is closed and bounded, there is a pair M_1, N_1 for which this value is actually attained. That is, we can obtain that

$$M_1AN_1 = \begin{bmatrix} d_1 & X \\ \\ Y & A' \end{bmatrix},$$

where A' is an $(m-1) \times (n-1)$-matrix. Now we claim that $X = 0$ and $Y = 0$ are *zero* rows and columns. Indeed, if X were non-trivial, the first row ρ_1 of M_1AN_1 would have length $d > d_1$. Then we could multiply on the right by the reflection H which takes ρ_1 into $[d, 0, \ldots, 0]$ and create a value $\alpha(M, N) = d > d_1$. Similarly $Y = 0$. Obviously, none of the entries of A' can exceed d_1 in absolute value (otherwise *it* could be permuted to the upper left), and this is true for all the possible forms of A'. We are finished by induction.

The real numbers $d_1 \geq, \ldots, \geq d_r > 0$ are known as the *singular values* of A.

UNIQUENESS: The $n \times n$ matrix $B = A^TA$ has a very simple effect on the columns $u_1, \cdots, u_n$ of N, namely, $Bu_i = \mu_i u_i$, where $\mu_i = d_i^2$ for $i \leq r$ and 0 beyond. Indeed, $N^TBN = (MAN)^TMAN = \Delta$ is a diagonal matrix with diagonal entries μ_i as described. Now the identity $BN = N\Delta$ establishes our claim.

To prove *uniqueness* of the singular values d_i it clearly suffices to characterize the μ_i as being the only numbers such that $(B - \mu I)u = 0$ for some $u \neq 0$. But for $u = \sum a_i u_i$, we get $(B - \mu I)u = \sum a_i(\mu_i - \mu)u_i$, which is never 0, unless μ is one of the μ_i.

More geometrically, the d_i can also be retrieved from the image under A of the appropriate unit sphere.

COROLLARY: Every symmetric $n \times n$ real matrix A has an eigenline.

Proof: Let $u \neq 0$ be one of the columns of N, so that $A^2u = A^TAu = \mu u$, as above. Put $\mu = \lambda^2$. Then u is annihilated by $A^2 - \mu I = (A + \lambda I)(A - \lambda I)$. If $(A - \lambda I)u = v \neq 0$, then v generates such a line; if $v = 0$ then u does.

For symmetric A it is trivial to show that the orthocomplement of any invariant subspace is itself invariant. Hence, by induction, the Corollary yields a set of n mutually orthogonal eigenlines (this is the famous "Spectral Theorem"). Moreover, if B is symmetric and commutes with A, it can be restricted to $\ker(A - \lambda I) \neq 0$; therefore the two matrices have a *common* eigenline, hence — by induction — a complete orthogonal set of such.

(All arguments on this page go through without a hitch for *complex* matrices if one changes "orthogonal" and "symmetric" to "unitary" and "hermitian", respectively, and replaces the transpose A^T by its complex conjugate $\bar{A}^T$. Writing a complex matrix as $C = A + iB$ with A, B hermitian, we again get a spectral theorem for C whenever A and B commute.)

7. Invariant Factors. Here is another variation on the "MAN" theme introduced in §5, this time applied to matrices with integer entries. Note that a strongly regular matrix M with entries in $\mathbf{Z}$ (the integers) is invertible over $\mathbf{Z}$, that is: M^{-1} also has integer entries (it suffices to check this for Gaussian matrices and permutation matrices, where it is obvious).

THEOREM: Let A be an $m \times n$ matrix over $\mathbf{Z}$. Then there exist strongly regular matrices M and N such that

$$MAN = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \qquad \text{where} \qquad D = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_r \end{bmatrix} \qquad \text{with} \qquad d_i \mid d_{i+1} \neq 0.$$

Proof: Let $\mathcal{S}$ be the set of all non-zero entries of MAN as M and N range over all strongly regular matrices of the appropriate sizes. Take $d_1 \in \mathcal{S}$ with $|d_1|$ minimal. By definition, we then have

$$M_1AN_1 = \begin{bmatrix} d_1 & X \\ Y & A' \end{bmatrix},$$

where X is a row, Y is a column, and A' is an $(m-1) \times (n-1)$-matrix. We claim that $X, Y \equiv 0$ modulo d_1.

Indeed, if y is any non-zero entry of Y (say in row μ), we may write it as $y = qd_1 + r$, with $|r| < |d_1|$. Multiplying M_1AN_1 on the left by the Gaussian matrix $I - qE_{\mu,1}$, we obtain the entry r in the place of y, contradicting the minimality of $|d_1|$, unless $r = 0$ as claimed. Hence we can make $Y = 0$ by such Gaussian multiplications, and similarly (by right multiplications) $X = 0$.

Assuming this done, we conclude that all entries of A' are divisible by d_1, because any one of them can be made to appear in the first column by a suitable addition of columns (i.e., Gaussian multiplication on the right), thus playing the role of the y in the argument above. The proof is finished by induction.

NOTE: The integers $d_1, \dots, d_r$ are known as the *invariant factors* of A. Their uniqueness can be proved via determinants of submatrices of A as follows. For every $\nu \leq \min(m, n)$, let $E_\nu(A) \subseteq \mathbf{Z}$ be the additive group generated by all $\nu \times \nu$ subdeterminants of A. Convince yourself that $E_\nu(A)$ remains unchanged by left or right multiplication of A by strongly regular matrices. Hence $E_\nu(A) = E_\nu(MAN) = (d_1 \cdots d_\nu)\mathbf{Z}$.

Congruence modulo a matrix.

An $m \times n$ integer matrix A can be used to define a congruence relation on the m-fold Cartesian product $\mathbf{Z}^m$ as follows: given two columns C_1 and C_2 in $\mathbf{Z}^m$, we write $C_1 \equiv C_2 \pmod{A}$ if $C_1 - C_2 = AX$ for suitable $X \in \mathbf{Z}^n$. A careful imitation of the proof given for the case $m = n = 1$ shows that this relation is compatible with addition (and "scalar multiplication" by individual integers). Hence the congruence classes form an additive group, denoted by $\mathbf{Z}^m/A\mathbf{Z}^n$.

What happens if A is multiplied on the left by an invertible matrix M? Well, $C_1 - C_2 = AX \iff MC_1 - MC_2 = MAX$, in other words, left multiplication by M changes a congruence class modulo A into a congruence class modulo MA, thus giving a homomorphism $\mathbf{Z}^m/A\mathbf{Z}^n \longrightarrow \mathbf{Z}^m/MA\mathbf{Z}^n$. Since M^{-1} reverses this map, it is an isomorphism.

What happens if A is multiplied on the right by an invertible matrix N? Nothing: $C_1 - C_2 = AX \iff C_1 - C_2 = ANX'$, because any X can be rewritten as NX', with $X' = N^{-1}X$.

COROLLARY: Let A be an $m \times n$ matrix over $\mathbf{Z}$ with invariant factors $d_1, \dots, d_r$. Then $\mathbf{Z}^m/A\mathbf{Z}^n$ is isomorphic to the direct product $\mathbf{Z}/d_1\mathbf{Z} \times \cdots \times \mathbf{Z}/d_r\mathbf{Z} \times \mathbf{Z}^{m-r}$.

Proof: Let M and N be as in the theorem, and put $A^* = MAN$. By the preceding discussion, M induces an isomorphism $\mathbf{Z}^m/A\mathbf{Z}^n \longrightarrow \mathbf{Z}^m/A^*\mathbf{Z}^n$.

8. Finite Abelian Groups. We start by characterizing subgroups of the additive group $\mathbf{Z}^m$.

LEMMA: Every subgroup of $\mathbf{Z}^m$ is of the form $A\mathbf{Z}^m$, where A is an $m \times m$ matrix.

Proof: If S is the given subgroup, we need to find m columns $C_1, \ldots, C_m$ in S such that every element of S can be expressed as $x_1 C_1 + \cdots + x_m C_m$ with $x_i \in \mathbf{Z}$.

Let $\lambda : S \longrightarrow \mathbf{Z}$ be the projection on the last component. Since $\ker(\lambda)$ can be identified with a subgroup $S' \subseteq \mathbf{Z}^{m-1}$, induction gives us columns $C_1, \ldots, C_{m-1}$ which generate S'. On the other hand, $\mathrm{im}(\lambda) = \lambda(S)$ is a subgroup of $\mathbf{Z}$ and therefore equal to $d\mathbf{Z}$ for some integer d. Pick $C_m \in S$ such that $\lambda(C_m) = d$.

Now for any $C \in S$, we have $\lambda(C) = kd$, and hence $C - kC_m \in S'$, with $k \in \mathbf{Z}$. Therefore $C - kC_m = x_1 C_1 + \cdots + x_{m-1} C_{m-1}$ with suitable $x_i \in \mathbf{Z}$, as desired.

THEOREM: Every finite abelian group is isomorphic to a direct product of cyclic groups.

Proof: If H is the finite group in question, it is clearly possible to find finitely many elements $h_1, \ldots, h_m$ which generate H. This gives a surjective homomorphism $\psi : \mathbf{Z}^m \longrightarrow H$ by

$$\psi(x_1, \ldots, x_m) = x_1 h_1 + \cdots + x_m h_m,$$

with H written additively. By the lemma, $\ker(\psi) = A\mathbf{Z}^m$ for a suitable matrix A, and by the corollary on the last page it follows (via First Iso. Thm.) that H is isomorphic to $\mathbf{Z}/d_1\mathbf{Z} \times \cdots \times \mathbf{Z}/d_m\mathbf{Z}$. (If you were too generous with your generators, some of the d_i will equal 1 and may be omitted in this decomposition.)

Two Further Decompositions.

Exercise 1: Let H be as above. For any integer n, define the subgroup $H(n) = \{x \in H \mid nx = 0\}$. Supposing that $n = dk$ with $(d, k) = 1$, show that

(a) $H(d) \cap H(k)$ is trivial $(= 0)$, and

(b) $H(n) = H(d) \times H(k)$ is a direct product.

For suitable n, we have $H(n) = H$. Therefore the prime factorization $n = p_1^{s_1} \cdots p_t^{s_t}$ and repeated application of this exercise yield the "primary decomposition"

$$H = H(p_1^{s_1}) \times \cdots \times H(p_t^{s_t}).$$

The components of this decomposition are *unique*. The first one, for instance, consists of all elements of H whose order is a power of p_1.

Exercise 2: With H still written additively, assume that $H = H(p^s)$ for some prime p. Applying the theorem, suppose you get

$$H = \mathbf{Z}/p^{\nu_1}\mathbf{Z} \times \cdots \times \mathbf{Z}/p^{\nu_r}\mathbf{Z},$$

with $\nu_1 \geq \cdots \geq \nu_r > 0$. Then we say that H is of type $(p^{\nu_1}, \cdots, p^{\nu_r})$. Prove:

(a) H is of type $(p^{\nu_1}, \cdots, p^{\nu_r}) \iff H(p)$ is of type $(p, \cdots, p)$ with r terms and pH is of type $(p^{\nu_1 - 1}, \cdots, p^{\nu_z - 1})$, where $z \leq r$ is the largest index with $\nu_z > 1$.

(b) Another abelian group K is isomorphic to H if and only if it is p-primary of the same type.

A convenient way of checking wheher two arbitrary finite abelian groups are isomorphic is to chop them up into primary components and then look at the type of the latter.