Math 500: Markov Processes, Decisions, and Evolution: Class #2

A Brief Review of Basic Linear Algebra (continued)

The standard definition of a nonsingular $(n \times n)$ matrix A is that there exists a matrix denoted A^{-1} such that $A^{-1}A = AA^{-1} = I$ (the identity matrix). It is not difficult to establish (you are encouraged to do it) the following:

Theorem: the following statements are equivalent

(1) A is nonsingular;

(2) A is row equivalent to the identity I (that is you can go from A to I, and conversely, by row operations);

(3) For any b, the equation $Ax = b$ has exactly one solution.

With the proper definitions and further developments, we will be able to add (at least) two statements to the above two:

(4) A has rank n (n being the common number of rows and columns of A);

(5) the determinant (det A) of A is non-zero.

To understand statement (4) more fully, let us discuss vector spaces and linear maps. To define a vector space we need a set **V** (of vectors) and a field **F** (of scalars, typically the real **R** or the complex **C**). We also need two operations:

Addition: $\forall u, v \in \mathbf{V}$, $u + v \in \mathbf{V}$;

Scalar Multiplication: $\forall u \in \mathbf{V}$, $\forall a \in \mathbf{F}$, $au \in \mathbf{V}$.

The two operations must satisfy several axioms (commutativity, associativity, additive identity, additive inverse, multiplicative identity, distributivity) for **V** to deserve the title of vector space over the field **F**.

A typical example for **V** is the set of n-tuples $(x_1,..,x_i,..,x_n) \in \mathbf{R}^n$ with operations

$(x_1,..,x_i,..,x_n) + (y_1,..,y_i,..,y_n) = (x_1 + y_1,..,x_i + y_i,..,x_n + y_n)$ and

$a(x_1,..,x_i,..,x_n) = (ax_1,..,ax_i,..,ax_n)$, $\forall a \in \mathbf{R}$.

A fundamental concept in the study of vector spaces is that of linear independence. A (sub)set of vectors $\mathbf{U} = \{u_1,..,u_i,..u_m\} \subset \mathbf{V}$ is *linearly independent* if the equality $a_1 u_1 + .. + a_i u_i + .. + a_m u_m = 0$ requires that $a_1 = .. = a_i = .. = a_m = 0$. Another important concept is that of *span*: the span of **U** is the set of all linear combinations of the form $a_1 u_1 + .. + a_i u_i + .. + a_m u_m$ with all u_i's in **U**. If a set **U** is linearly independent and spans all of **V**, it is called a *basis* of **V**. An important result is that if a vector space admits a basis made up of n vectors, then any basis of **V** is made up of n vectors and we say that **V** has dimension n.

The next concept is that of linear map T:**V** $\longrightarrow$ **V**, from a vector space **V** into itself, which must satisfy (with Tu denoting the image by T of u):

$$T(u+v) = Tu + Tv$$

$$T(\mathrm{a}u) = \mathrm{a}Tu$$

Once a basis (say **B**=$\{e_1,..,e_j,..,e_n\}$) has been chosen for the vector space **V**, the vectors $Te_1,..,Te_j,..,Te_n$ are of special interest. Since **B** is a basis, any vector u can be written (uniquely) as a linear combination $u = \mathrm{x}_1 e_1 + .. + \mathrm{x}_j e_j + .. + \mathrm{x}_n e_n$ (x_j is called the jth component of u) and u can be alternatively written as a $n \times 1$ (column) matrix $\mathrm{u} = \begin{pmatrix} \mathrm{x}_1 \\ \mathrm{x}_j \\ \mathrm{x}_n \end{pmatrix}$ or as a $1 \times n$ (row) matrix $\mathrm{u}^T = (\,\mathrm{x}_1,..,\mathrm{x}_j,..,\mathrm{x}_n\,)$.

In particular, each Te_j can be written $Te_j = \mathrm{a}_{1j}e_1 + .. + \mathrm{a}_{ij}e_i + .. + \mathrm{a}_{nj}e_n$. Similarly, Tu can be written $\mathrm{x}_1 Te_1 + .. + \mathrm{x}_j Te_j + .. + \mathrm{x}_n Te_n$. So, if we write the $n \times n$ matrix A whose columns are the components of the Te_j's: $\mathrm{A} = \begin{pmatrix} \mathrm{a}_{11}.. & \mathrm{a}_{1j}.. & \mathrm{a}_{1n} \\ \mathrm{a}_{i1..} & \mathrm{a}_{ij}.. & \mathrm{a}_{in} \\ \mathrm{a}_{n1}.. & \mathrm{a}_{nj}.. & \mathrm{a}_{nn} \end{pmatrix}$ then the vector Tu has components given by the product Au.

Now, the set $\mathbf{U} = \{Te_1,..,Te_i,..,Te_n\}$ may either span **V** or some (strict) subset of **V** (called the *range* of T) that is a *subspace* because it is a vector space in its own right. That subspace has its own dimension (less than or equal to n) which is also called the *rank* of the matrix A (or the linear map T). Another subspace of interest is the *null* (or *kernel*) of T which is the set of vectors that T maps into the zero vector, another subspace in its own right. A standard result is that the dimension of the span of T (the rank) and the dimension of the null of T add up to the dimension n of **V**.

If the rank of T is equal to n, it means that the above set **U** spans **V**. Since it is made up of n vectors, it is a basis of **V** (you can always extract a basis from a spanning set but that basis must have n vectors). So, any vector v of components b_j in the basis **B** can be *uniquely* written $v = \mathrm{x}_1 Te_1 + .. + \mathrm{x}_j Te_j + .. + \mathrm{x}_n Te_n$. This means that $\mathrm{b} = \mathrm{Ax}$ where x is the unique $n \times 1$ (column) matrix of elements x_j. So, if A has rank n, the solution of $\mathrm{Ax} = \mathrm{b}$ always exists and is always unique.

The next important concept is that of *dot* or *inner* product of two vectors. Here we need to be specific about what field (**R** or **C**) we assume with our vector space. In the real case, the dot product of u by v, denoted $u.v$ (or <u,v>) is simply

$$u.v = \mathrm{u}^T\mathrm{v} = \mathrm{x}_1\mathrm{y}_1 + .. + \mathrm{x}_i\mathrm{y}_i + .. + \mathrm{x}_n\mathrm{y}_n$$

where u^T is the row matrix of components (the x_i's) of u on the chosen basis and v is the column matrix of components (the y_i's) of v in that same basis.

If our field is **C**, we modify this into

$$u.v = u^* v = \bar{x}_1 y_1 + .. + \bar{x}_i y_i + .. + \bar{x}_n y_n$$

where $\bar{x}_i$ is the conjugate of the complex number x_i.

This guarantees that, in either case, the dot product satisfies the necessary axioms of positivity, definitness, additivity, homogeneity, and (conjugate) symmetry. One can then define the norm of a vector by $\|u\| = \sqrt{u.u}$. A vector is called normal if $\|u\| = 1$. And two vectors u and v are said to be orthogonal if $u.v = 0$.

It is not difficult to show how one can construct from any basis of **V** another basis made up of normal vectors that are all orthogonal to each other. The process is called Gram-Schmidt orthonormalization and the result is called an orthonormal basis.

Homework

1. Argue that any matrix A has at most one inverse.

2. Look into the literature to justify (and explain) the following statement: "An automorphism is a particular case of isomorphism."

3. Prove that the null of a linear map is a vector space.

4. Prove that a set of non-zero orthogonal vectors is linearly independent.