University of Regina
Department of Mathematics and Statistics

A LINEAR ALGEBRA DICTIONARY FOR STUDENTS OF MATHEMATICS 122

(to be used as a companion to the lectures and the readings in the text)

Author: D Farenick

Last Update: November 14, 2006

- *Linear Equation.* A linear equation in p variables $x_1, \ldots, x_p$ is an equation of the form

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_p x_p = b ,$$

where $\alpha_1, \ldots, \alpha_p, b$ are fixed real numbers. For example, $x_1 - x_2 + 5x_3 = \sqrt{2}$ is a linear equation, but $x_1^3 + \sqrt{x_2 + 1} = 1$ is not.

- *Systems of Linear Equations.* A system of m linear equations in p variables $x_1, \ldots, x_p$ is a collection of m linear equations considered simultaneously:

$$\begin{array}{ccccccccc} \alpha_{11}x_1 & + & \alpha_{12}x_2 & + & \cdots & + & \alpha_{1p}x_p & = & b_1 \\ \alpha_{21}x_1 & + & \alpha_{22}x_2 & + & \cdots & + & \alpha_{2p}x_p & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \alpha_{m1}x_1 & + & \alpha_{m2}x_2 & + & \cdots & + & \alpha_{mp}x_p & = & b_m . \end{array} \tag{1}$$

The system (1) of linear equations is said to be:

(a) *consistent*, if there are real numbers $x_1, \ldots, x_p$ that satisfy all m equations (1) simultaneously; or

(b) *inconsistent*, if no real numbers $x_1, \ldots, x_p$ that satisfy all m equations (1) simultaneously.

- *Systems of Homogeneous Linear Equations.* A system of m linear equations in p variables $x_1, \ldots, x_p$ is homogeneous if it is of the form

$$\begin{array}{ccccccccc} \alpha_{11}x_1 & + & \alpha_{12}x_2 & + & \cdots & + & \alpha_{1p}x_p & = & 0 \\ \alpha_{21}x_1 & + & \alpha_{22}x_2 & + & \cdots & + & \alpha_{2p}x_p & = & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \alpha_{m1}x_1 & + & \alpha_{m2}x_2 & + & \cdots & + & \alpha_{mp}x_p & = & 0 . \end{array} \tag{2}$$

Observe that if $x_1 = x_2 = \cdots = x_p = 0$, then $x_1, \ldots, x_p$ satisfy all m equations (2) simultaneously. That is, every system of homogeneous equations is consistent.

- *Vectors and* $\mathbb{R}^p$. Let $\mathbb{R}$ denote the set of all real numbers. By $\mathbb{R}^p$ is meant the set of all p-tuples of real numbers, written in column form. That is, $x \in \mathbb{R}^p$ if and only if

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}, \quad \text{where } x_1, x_2, \cdots, x_p \in \mathbb{R}.$$

The elements of $\mathbb{R}^p$ are called vectors.

Geometrically, a vector $x = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix} \in \mathbb{R}^p$ can be viewed as a directed line segment form the origin ($0 \in \mathbb{R}^p$) to the point with Cartesian coordinates $(x_1, \ldots, x_p)$. Hence, vectors are distinguished from points in that they have both direction and magnitude. Nevertheless, we still consider that a vector $x \in \mathbb{R}^p$ determines a point in space: namely, the terminal point of the vector x.

- *Linear Equations in* $\mathbb{R}^2$ *and* $\mathbb{R}^3$. A linear equation in two variables x_1 and x_2 determines a line in $\mathbb{R}^2$. That is, the set of $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$ that satisfy the linear equation

$$\alpha_1 x_1 + \alpha_2 x_2 = b,$$

where α_1, α_2, b are fixed real numbers with at least one nonzero is a line. Indeed, if one wished to write the equation of the line in the form $y = mx + b$, then, assuming $\alpha_2 \neq 0$, we obtain $x_2 = \frac{-\alpha_1}{\alpha_2} x_1 + \frac{b}{\alpha_2}$. However, this is expression is not very useful in linear algebra and we prefer to express lines by either the linear equation $\alpha_1 x_1 + \alpha_2 x_2 = b$ or by the parametric form

$$x = v_0 + tv, \quad t \in \mathbb{R},$$

where $v_0 \in \mathbb{R}^2$ is a vector whose terminal point is on the line (and whose initial point is the origin) and $v \in \mathbb{R}^2$ determines the direction of the line.

A linear equation in three variables x_1, x_2, and x_3 determines a plane in $\mathbb{R}^3$. That is, the set of $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$ that satisfy the linear equation

$$\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = b,$$

where $\alpha_1, \alpha_2, \alpha_3, b$ are fixed real numbers with at least one nonzero, is a plane. As with lines, one can express the equation of a plane in parametric form:

$$x = v_0 + sv_1 + tv_2,, \quad s, t \in \mathbb{R},$$

where $v_0 \in \mathbb{R}^3$ is a vector whose terminal point is on the plane (and whose initial point is the origin) and $v_1, v_2 \in \mathbb{R}^2$ determine the orientation of the plane.

- *Systems of Linear Equations in* $\mathbb{R}^2$ *and* $\mathbb{R}^3$. The solutions $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$ to a system of m linear equations in two variables x_1 and x_2 are all points x (*ie.*, the terminal point of the vector x)

that lie on each of the m lines determined by the m linear equations. In other words, a system of m linear equations in two variables describes the intersection of m lines in the plane.

The solutions $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$ to a system of m linear equations in three variables x_1, x_2, and x_3 are all points x (*ie.*, the terminal point of the vector x) that lie on each of the m planes determined by the m linear equations. In other words, a system of m linear equations in three variables describes the intersection of m planes in $\mathbb{R}^3$.

- *Matrices and $M_{m,p}(\mathbb{R})$.* An $m \times p$ array of real numbers α_{ij} is called an $m \times p$ matrix. The subscript ij of α_{ij} indicates that the real number α_{ij} is in the i-th row and j-th column of the array. The set of $m \times p$ matrices is denoted by $M_{m,p}(\mathbb{R})$ and an arbitrary matrix $A \in M_{m,p}(\mathbb{R})$ is denoted by

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1p} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mp} \end{bmatrix}. \tag{3}$$

- *Matrix-Vector Product.* If $A \in M_{m,p}$ and $x \in \mathbb{R}^p$—that is, A is an $m \times p$ matrix and x is a vector in $\mathbb{R}^p$—then Ax denotes the vector in $\mathbb{R}^m$ defined by

$$Ax = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1p} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mp} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{p} \alpha_{1j} x_j \\ \sum_{j=1}^{p} \alpha_{2j} x_j \\ \vdots \\ \sum_{j=1}^{p} \alpha_{mj} x_j \end{bmatrix}.$$

If $A \in M_{m,p}(\mathbb{R})$ is as in (3), and if

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} \in \mathbb{R}^p \quad \text{and} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m,$$

then $x_1, \ldots, x_p$ satisfy the system (1) if and only if $Ax = b$.

- *Linear Combinations.* If v_1, v_2, $\ldots$, v_p are vectors in $\mathbb{R}^m$, then a linear combination of these vectors is an expression of the form

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_p v_p, \tag{4}$$

for some p real numbers $\alpha_1, \ldots, \alpha_p$. For example, $v_1 - v_2 + \sqrt{3} v_3$ is a linear combination of vectors $v_1, v_2, v_3 \in \mathbb{R}^m$. Because, in $\mathbb{R}^m$, one can add vectors and multiply vectors by scalars (real numbers), the result of a linear combination such as (4) is another vector in $\mathbb{R}^m$.

- *Span.* If $v_1, v_2, \ldots, v_p$ are vectors in $\mathbb{R}^m$, then the span of these vectors is the set denoted by $\text{Span}\,\{v_1, \ldots, v_p\}$ and defined by

$$\text{Span}\,\{v_1, \ldots, v_p\} \;=\; \left\{\alpha_1 v_1 \,+\, \alpha_2 v_2 \,+\, \cdots \,+\, \alpha_p v_p \,,|\, \alpha_1, \ldots, \alpha_p \in \mathbb{R}\right\}\,. \tag{5}$$

  In other words, $\text{Span}\,\{v_1, \ldots, v_p\}$ is the set of all linear combinations of the vectors $v_1, \ldots, v_p \in \mathbb{R}^m$. Note that each v_j belongs to $\text{Span}\,\{v_1, \ldots, v_p\}$, as does the zero vector $0 \in \mathbb{R}^m$.

  How can one determine whether a given vector $b \in \mathbb{R}^m$ belongs to the spanning set $\text{Span}\,\{v_1, \ldots, v_p\}$? There is an extremely useful <u>criterion</u>:

$$b \in \text{Span}\,\{v_1, \ldots, v_p\} \text{ if and only if } Ax = b \text{ is consistent!}$$

  (In the statement above, A is the $m \times p$ matrix whose j-th column is the vector v_j.) Any solution $x \in \mathbb{R}^p$ to $Ax = b$ determines b as a linear combination of $v_1, \ldots, v_p$ (the columns of A). For example, if $v_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $v_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, and $v_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, then A would be

$$A \;=\; [v_1,\; v_2,\; v_3] \;=\; \begin{bmatrix} -1 & -2 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

  Further, if $b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, then

$$b \in \text{Span}\,\{v_1, v_2, v_3\}$$

  because the equation $Ax = b$ admits a solution $x = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$. The entries of x determine an explicit manner for b to be expressed as a linear combination of v_1, v_2, v_3:

$$b \;=\; -v_1 \,+\, v_3\,.$$

  That is,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = -\begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

- *Linear Independence/Dependence.* If $v_1, v_2, \ldots, v_p$ are vectors in $\mathbb{R}^m$, then these vectors are said to be linearly independent if the only $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$ that satisfy the equation

$$\alpha_1 v_1 \,+\, \alpha_2 v_2 \,+\, \cdots \,+\, \alpha_p v_p \;=\; 0 \tag{6}$$

  are $\alpha_1 = 0$, $\alpha_2 = 0$, $\ldots$, and $\alpha_p = 0$.

  If the vectors $v_1, v_2, \ldots, v_p$ are not linearly independent, then they are said to be linearly dependent.

  How can one determine whether $v_1, v_2, \ldots, v_p$ are linearly independent vectors? There is an extremely useful <u>criterion</u>:

$$v_1, \ldots, v_p \text{ are linearly independent if and only if } Ax = 0 \text{ has a unique solution } x\,!$$

  (In the statement above, A is the $m \times p$ matrix whose j-th column is the vector v_j; the solution x would necessarily be the trivial solution $x = 0 \in \mathbb{R}^p$.)

  What do these concepts mean? The answer lies in the following two (logically equivalent) theorems.

THEOREM 1. *$v_1, \ldots, v_p \in \mathbb{R}^m$ are linearly independent if and only if no vector v_j is a linear combination of the other vectors $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_p$.*

THEOREM 2. *$v_1, \ldots, v_p \in \mathbb{R}^m$ are linearly dependent if and only if at least one vector v_j is a linear combination of the other vectors $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_p$.*

- *Linear Transformations.* A function $T : \mathbb{R}^p \to \mathbb{R}^m$ is a linear transformation if the following two conditions are satisfied:

  (i) $T(u + v) = T(u) + T(v)$, for all $u, v \in \mathbb{R}^p$;
  (ii) $T(\alpha u) = \alpha T(u)$, for all $u \in \mathbb{R}^p$ and $\alpha \in \mathbb{R}$.

  Every $m \times p$ matrix T is a linear transformation $\mathbb{R}^p \to \mathbb{R}^m$ and, conversely, every linear transformation $T : \mathbb{R}^p \to \mathbb{R}^m$ is a matrix whose j-th column is the vector $T(e_j) \in \mathbb{R}^m$.

  What does this concept mean ? If T is an $m \times p$ matrix, and if

$$x = \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_p \end{bmatrix} \in \mathbb{R}^p ,$$

  then

$$Tx = \begin{bmatrix} \sum_{j=1}^{p} t_{1j} x_j \\ \vdots \\ \vdots \\ \sum_{j=1}^{p} t_{mj} x_j \end{bmatrix} \in \mathbb{R}^m .$$

  Each component of the vector Tx is a linear function (equation) in the “variables" $x_1, \ldots, x_p$. Thus,

  PROPOSITION. *A function $T : \mathbb{R}^p \to \mathbb{R}^m$ is a linear transformation if and only if each of the m components of the vector $Tx \in \mathbb{R}^m$ is a linear equation in the components $x_1, \ldots, x_p$ of the vector $x \in \mathbb{R}^p$.*

- *Matrix Multiplication.* If A is an $m \times q$ matrix and B is a $q \times p$ matrix, then one can form the product AB to obtain a $m \times p$ matrix whose (i, j)-entry is defined to $\sum_{k=1}^{q} \alpha_{ik} \beta_{kj}$, where α_{st} and $\beta_{\mu\nu}$ denote the entries of A and B respectively.

  If the columns of B are given by the vectors $b_1, \ldots, b_p \in \mathbb{R}^q$, then it is convenient to view the product AB as the result of p matrix-vector products: namely, the j-th column of AB is Ab_j. We denote this as follows:

$$AB = A \begin{bmatrix} b_1 & b_2 & \ldots & b_p \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \ldots & Ab_p \end{bmatrix} .$$

  Square matrices A and B are said to commute if $AB = BA$. However, for most $A, B \in M_p(\mathbb{R})$ we have $AB \neq BA$. Hence, matrix multiplication is noncommutative.

- *Elementary Matrices.* A matrix $E \in M_p(\mathbb{R})$ is an elementary matrix if E is obtained from the identity matrix I by applying one elementary row operation to I. Recall that there are three elementary row operations:

  (i) αR_i (multiply the i-th row by $\alpha \in \mathbb{R}$);

  (ii) $R_i \leftrightarrow R_k$ (exchange rows i and k); and

  (iii) $\alpha R_i + R_k \longrightarrow R_k$ (multiply the i-th row by α, add this to the k-th row and put the result back into the k-th row).

  For example,

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$$

  is an elementary matrix (obtained from the row operation $(-2)R_2 + R_3 \longrightarrow R_3$ applied to I).

- *The Transpose.* If $A \in M_{m,p}(\mathbb{R})$, then the transpose of A is the matrix $A^T \in M_{p,m}$ whose from A by making the rows of A become the columns of A^T. More precisely, the (i, j)-entry of A^T is the (j, i)-entry of A, for each $i = 1, \ldots, m$ and $j = 1, \ldots, p$. Note that $A^T A$ and AA^T are square matrices. An important identity relates to products: $(AB)^T = B^T A^T$.

- *The Inverse of a Square Matrix.* A matrix $A \in M_p(\mathbb{R})$ is invertible if there is a matrix $C \in M_p(\mathbb{R})$ such that $AC = CA = I$. If such a matrix C exists—that is, if A is invertible—then C is necessarily unique and we use the notation A^{-1} to denote C. Thus,

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

  Use: if A is invertible, then the matrix equation $Ax = b$ has a (unique) solution x given by $x = A^{-1}b$.

  An important identity relates to products of invertible matrices: $(AB)^{-1} = B^{-1}A^{-1}$. Furthermore, if A is invertible, then so is A^T.

  To compute A^{-1}, apply elementary row operations to reduce the augmented matrix $[A : I]$ to the form $[I : C]$. The matrix C the inverse A^{-1} of A. This means, therefore, that I is a product $E_q E_{q-1} \cdots E_2 E_1$ of q elementary matrices E_j corresponding to the q elementary row operations that were used to reduce $[A : I]$ to $[I : C]$.

  If A is invertible, then $Ax = 0$ if and only if $x = A^{-1}0 = 0$. Therefore, the columns of an invertible matrix A are linearly independent. Likewise, the columns of A^T must be linearly independent. Hence, a matrix A is invertible only if the rows and columns of A are linearly independent. This fact can be used to quickly identify non-invertible matrices.

  An important theorem linking these notions is as follows.

  THEOREM. *The following statements are logically equivalent for a matrix $A \in M_p(\mathbb{R})$:*

  (a) the columns of A are linearly independent;

  (b) the columns of A span $\mathbb{R}^p$;

  (c) A is an invertible matrix.

- *Vector Spaces.* A vector space (over the real numbers $\mathbb{R}$) is a nonempty set V endowed with operations of addition and scalar multiplication such that:

  (a) the operation of addition satisfies

    (i) $v_1 + v_2 = v_2 + v_1$, for all $v_1, v_2 \in V$,

    (ii) $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$, for all $v_1, v_2, v_3 \in V$,

    (iii) there is a unique element $0 \in V$ such that $v + 0 = 0 + v = v$, for all $v \in V$, and

    (iv) for every $v \in V$ there is a unique element $(-v) \in V$ such that $v + (-v) = 0$;

  (b) and the operation of scalar multiplication satisfies, for all $\alpha, \alpha_1, \alpha_2 \in \mathbb{R}$ and $v, v_1, v_2 \in V$,

    (i) $(\alpha_1\alpha_2)v = \alpha_1(\alpha_2 v)$,

    (ii) $\alpha v = 0 \in V$ whenever $\alpha = 0 \in \mathbb{R}$ or $v = 0 \in V$,

    (iii) $1\,v = v$, where $1 \in \mathbb{R}$,

    (iv) $\alpha(v_1 + v_2) = \alpha v_1 + \alpha v_2$,

    (v) $(\alpha_1 + \alpha_2)v = \alpha_1 v + \alpha_2 v$.

  Some familiar vector spaces are:

  (a) our old friend $\mathbb{R}^p$,

  (b) the space $M_{m,p}(\mathbb{R})$ of $m \times p$ matrices,

  (c) the space $\mathfrak{F}(\mathbb{R})$ of all functions $f : \mathbb{R} \to \mathbb{R}$,

  (d) the space $\mathbb{R}[x]$ of all polynomials (with real coefficients), and

  (e) the space $\mathbb{R}_{(p)}[x]$ of all polynomials of degree at most p.

- *Subspaces.* A subspace W of a vector space V is a subset $W \subseteq V$ such that

  (a) $w_1 + w_2 \in W$, for all $w_1, w_2 \in W$, and

  (b) $\alpha w \in W$, for every $\alpha \in \mathbb{R}$ and all $w \in W$.

  If $W \subseteq V$ is a subspace of a vector space V, then W itself is a vector space.

  In practice, if one is asked to show (or prove) that a certain subset W of a vector space V is a subspace, then one must verify that W satisfies the two conditions (a) and (b) above that define the notion of subspace.

  Examples of subspaces are:

  (a) $\mathbb{R}_{(p)}[x]$ is a subspace of $\mathbb{R}[x]$;

  (b) if $v_1, \ldots, v_p \in V$, then $\text{Span}\{v_1, \ldots, v_p\}$ is a subspace of V;

  (c) the subspaces of $\mathbb{R}^2$ are:

    (i) the zero subspace $\{0\}$ (where $0 \in \mathbb{R}^2$ is the origin);

    (ii) any line passing through the origin;

    (iii) $\mathbb{R}^2$ itself.

  (d) the space of all $A \in M_{m,p}(\mathbb{R})$ whose first column is zero is a subspace of $M_{m,p}(\mathbb{R})$.

- *Range and Kernel; Column Space and Null Space.* Two other examples of subspaces are important, and these arise as follows. Assume that U and V are vector spaces and that $T : U \to V$ is a linear transformation.

  The kernel of T is the subset $\ker T$ of U defined by: $\ker T \; = \; \{u \in U \mid Tu = 0\}$.

  The range of T is the subset Ran T of V defined by: $\operatorname{Ran} T \; = \; \{Tu \in V \mid u \in U\}$.

  THEOREM *If $T : U \to V$ is a linear transformation, then* $\ker T$ *is a subspace of U and* $\operatorname{Ran} T$ *is a subspace of V.*

  This theorem above can be proved by students of MATH 122. For example, to show that $\ker T$ is a subspace of U there are two conditions to verify:

  (a) that $u_1 + u_2 \in \ker T$, for all $u_1, u_2 \in \ker T$, and

  (b) that $\alpha u \in \ker T$, for every $\alpha \in \mathbb{R}$ and all $u \in \ker T$.

  To verify (a), let $u_1, u_2 \in \ker T$. To show that $u_1 + u_2 \in \ker T$ is to show that $T(u_1 + u_2) = 0$, since this is how the kernel of T is defined. To this end, compute: $T(u_1 + u_2) = Tu_1 + Tu_2 = 0 + 0 = 0$. This shows that $u_1 + u_2 \in \ker T$, for all $u_1, u_2 \in \ker T$. To verify (b), let $\alpha \in \mathbb{R}$ and $u \in \ker T$. Compute: $T(\alpha u) = \alpha T u = \alpha(0) = 0$. This proves that $\alpha u \in \ker T$, for every $\alpha \in \mathbb{R}$ and all $u \in \ker T$. (Note how the fact that T is a linear transformation was used in both (a) and (b).) Hence, we have proved that $\ker T$ is a subspace of U.

  If $U = \mathbb{R}^p$ and $V = \mathbb{R}^m$, then T is an $m \times p$ matrix and $\ker T$ is called the nullspace of T, denoted by Null T, and Ran T is called the column space of T, denoted by Col T.

- *Row Space.* Each row of an $m \times p$ matrix A can be considered as a "row" vector. Rather than writing a row as a vector in $\mathbb{R}^p$—which is a notation normally reserved for column vectors—let us denote the space of all row vectors by $(\mathbb{R}^p)^*$. The row space of A, denoted by Row A, is the subspace of $(\mathbb{R}^p)^*$ that is spanned by the rows of A.

- *Basis.* If W is a subspace of a vector space V, then vectors $w_1, \ldots, w_q \in W$ form a basis of W if

  (a) $w_1, \ldots, w_q$ are linearly independent, and

  (b) $W = \operatorname{Span}\{w_1, \ldots, w_q\}$.

  Examples of Bases for Vector Spaces are:

  (a) $e_1, \ldots, e_p$ form a basis of $\mathbb{R}^p$, where $e_j \in \mathbb{R}^p$ denotes the vector with 1 in position j and 0 elsewhere;

  (b) $E_{11}, \ldots, E_{1p}, E_{21}, \ldots, E_{2p}, \ldots, E_{m1}, \ldots, E_{mp}$ form a basis of $M_{m,p}(\mathbb{R})$, where each $E_{ij} \in M_{m,p}(\mathbb{R})$ is the matrix with (i, j)-entry 1 and zeros elsewhere in the matrix;

  (c) $1, x, x^2, x^3, x^4, \ldots$ form a basis of $\mathbb{R}[x]$;

  (d) $1, x, x^2, \ldots, x^p$ form a basis of $\mathbb{R}_{(p)}[x]$.

  The value of a basis is: if $w_1, \ldots, w_q$ are basis vectors for a subspace W, then every vector $w \in W$ has a unique representation as a linear combination of the basis vectors $w_1, \ldots, w_q$.

  Even more good things occur with bases. Let V be a p-dimensional vector space, and suppose that the vectors $v_1, \ldots, v_p \in V$ form a basis of V. Then each vector $v \in V$ is represented as

  $$v \; = \; \alpha_1 v \, + \, \alpha_2 v_2 \, + \, \cdots \, + \, \alpha_p v_p \, ,$$

where $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$ are uniquely determined by (*ie.*, identified with) v. These numbers $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$ are the "coordinates" of v with respect to the given basis. Therefore, we identify $v \in V$ with the coordinate vector $\mathfrak{v} \in \mathbb{R}^p$ defined by

$$\mathfrak{v} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix} \in \mathbb{R}^p .$$

For example, assume that $v_1 = \begin{bmatrix} -3 \\ 1 \\ -2 \end{bmatrix}$, $v_2 = \begin{bmatrix} 6 \\ \frac{3}{2} \\ \frac{7}{2} \end{bmatrix}$, and $v_3 = \begin{bmatrix} 15 \\ \frac{9}{2} \\ \frac{17}{2} \end{bmatrix}$ form a basis of $\mathbb{R}^3$ and consider the vector $v = \begin{bmatrix} -3 \\ \frac{1}{2} \\ -2 \end{bmatrix}$. To find the coordinates of v with respect to the basis vectors v_1, v_2, v_3 we must solve

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3$$

for $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$. To do so form the augmented matrix

$$\begin{bmatrix} -3 & -6 & -15 & : & -3 \\ 1 & 3/2 & 9/2 & : & 1/2 \\ -2 & -7/2 & -17/2 & : & -2 \end{bmatrix}$$

and use elementary row operations to reduce it to an echelon form:

$$\begin{bmatrix} -1 & -2 & -5 & : & -1 \\ 0 & -1 & -1 & : & 11 \\ 0 & 0 & -2 & : & 1 \end{bmatrix} .$$

Now back substitute and solve for each α_j to get $\alpha_1 = 1/2$, $\alpha_2 = 3/2$, and $\alpha_3 = -1/2$. Hence, the coordinate vector of v is the vector

$$\mathfrak{v} = \begin{bmatrix} 1/2 \\ 3/2 \\ -1/2 \end{bmatrix} .$$

- *Dimension.* The number of elements in any basis of a subspace W is called the dimension of W, denoted by $\dim W$.

  THEOREM *If* $A \in M_{m,p}(\mathbb{R})$, *then* $\dim(\text{Col } A) = \dim(\text{Row } A)$.

  We make the following definitions for $A \in M_{m,p}(\mathbb{R})$.

  (a) The rank of A is the number $\text{rank}(A)$ defined by: $\text{rank}(A) = \dim(\text{Col } A)$.

  (b) The nullity of A is the number $\text{nullity}(A)$ defined by: $\text{nullity}(A) = \dim(\text{Nul } A)$.

  RANK-PLUS-NULLITY THEOREM *If* $A \in M_{m,p}(\mathbb{R})$, *then* $p = \text{rank}(A) + \text{nullity}(A)$.

- *Finding the Bases for Null Spaces, Column Spaces, and Row Spaces.* We aim to find a basis for each of the subspaces Nul A, Col A, and Row A, given a matrix $A \in M_{m,p}(\mathbb{R})$.

  (a) *Null Spaces.* To compute a basis for Nul A, solve $Ax = 0$ by reducing the augmented matrix $[A : 0]$ to an echelon form. If there is a unique solution, then Nul $A = \{0\} \subset \mathbb{R}^p$. If there are infinitely many solutions, then identify the free variables, say $x_p, x_{p-1}, \ldots, x_q$, and express the general solution to $Ax = 0$ as a linear combination of vectors $v_p, v_{p-1}, \ldots, v_q \in \mathbb{R}^p$ using the free variables as coefficients. That is, $v \in$ Nul A if and only if

  $$v = x_p v_p + x_{p-1} v_{p-1} + \cdots + x_q v_q .$$

  The vectors $v_p, v_{p-1}, \ldots, v_q \in \mathbb{R}^p$ form a basis of Nul A.

  (b) *Column Spaces.* To compute a basis for Col A, reduce the matrix A to an echelon form B. In each row of B that has a leading nonzero entry, note the column in which the leading nonzero entry appears. The corresponding column in the original matrix A will be a basis vector for Col A. Doing this for every row of B that has a nonzero leading entry determines the basis vectors for Col A.

  (c) *Row Spaces.* To compute a basis for Row A, reduce the matrix A to an echelon form B. Each row of B that has a leading nonzero entry is a basis vector for Row A. Hence, the rows of B that have a nonzero leading entry are the basis vectors for Row A.

  REMARK. A basis for Col A consists of a subset of the columns of A. In contrast, a basis for Row A consists of a subset of the rows of an *echelon* form of A.

  ADDITIONAL REMARK. If $v_1, \ldots, v_p \in V$ and $W = \text{Span}\{v_1, \ldots, v_p\}$, then there is a subset of the vectors $v_1, \ldots, v_p$ that forms a basis for W.

- *The Matrix of a Linear Transformation.* Assume that $T : V \to W$ is a linear transformation. Let $v_1, \ldots, v_p \in V$ be basis vectors for V and $w_1, \ldots, w_m$ be basis vectors of W. The matrix of T with respect to these bases of V and W is the $m \times p$ matrix whose j-th column is given by the coordinates of Tv_j—that is, the j-th column is

  $$\begin{bmatrix} \tau_{1j} \\ \tau_{2j} \\ \vdots \\ \tau_{mj} \end{bmatrix} \in \mathbb{R}^m ,$$

  where $\tau_{1j}, \ldots, \tau_{mj} \in \mathbb{R}$ are the unique numbers for which

  $$Tv_j = \tau_{1j} w_j + \tau_{2j} w_j + \cdots + \tau_{mj} w_m .$$

- *Finding the Bases for the Kernel and Range of a Linear Transformation.* Assume that $T : V \to W$ is a linear transformation. To find a basis for ker T and ran T, the approach will depend on how much information about V, W, and T one has. An algortihmic approach—which is not necessarily the recommenced approach—is to (i) find bases for V and W, (ii) find the matrix representation of T with respect to these bases, (iii) find bases for the column and null space of the matrix representation, and (iv) use the coordinate vectors that form bases for the column and null space of the matrix to determine basis vectors for the kernel and range of the transformation.

- *Determinants.* The determinant is a number denoted by $\det A$ that is assigned to each $p \times p$ matrix A, for any p. The number $\det A$ is called the determinant of A.

  If $p = 1$ and $A = [a] \in M_1(\mathbb{R})$, then $\det A = a$.

  If $p = 2$ and $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{R})$, then $\det A = ad - bc$.

  If $p = 3$ and $A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \in M_3(\mathbb{R})$, then

  $$\det A \;=\; \alpha_{11}(-1)^{1+1}\det A_{11} \;+\; \alpha_{12}(-1)^{1+2}\det A_{12} \;+\; \alpha_{13}(-1)^{1+3}\det A_{13}\,,$$

  where A_{1j} denotes the 2×2 matrix that remains when row #1 and column #j of A are removed from A.

  If p is arbitrary and $A \in M_p(\mathbb{R})$, then

  $$\det A \;=\; \alpha_{11}(-1)^{1+1}\det A_{11} \;+\; \alpha_{12}(-1)^{1+2}\det A_{12} \;+\; \cdots \;+\; \alpha_{1p}(-1)^{1+p}\det A_{1p}\,,$$

  where A_{1j} is the $(p-1) \times (p-1)$ matrix that remains when row #1 and column #j of A are removed from A.

  The cofactors of $A \in M_p(\mathbb{R})$ are the numbers $C_{ij} = (-1)^{i+j}\det A_{ij}$, where A_{ij} is the $(p-1) \times (p-1)$ matrix that remains when row #i and column #j are removed from A.

  It is extremely useful that $\det A$ can be computed by a cofactor expansion along any of its rows or columns.

  THEOREM. *If $A \in M_p(A)$, then, for any i & j,*

  $$\begin{aligned} \det A \;&=\; \alpha_{i1}C_{i1} \;+\; \alpha_{i2}C_{i2} \;+\; \cdots \;+\; \alpha_{ip}C_{ip} \\ &=\; \alpha_{1j}C_{1j} \;+\; \alpha_{2j}C_{2j} \;+\; \cdots \;+\; \alpha_{pj}C_{pj}\,. \end{aligned}$$

  The effect of elementary row operations on the determinant are summarised as follows:

  (a) if A' is the matrix obtained from A by interchanging two rows of A, then $\det A' = (-1)\det A$;

  (b) if A' is the matrix obtained from A by multiplying row #i of A by $\alpha \in \mathbb{R}$, then $\det A' = \alpha \det A$; and

  (c) if A' is the matrix A except that row #j of A' is obtained from A by multiplying row #i of A by α and adding this to row #j of A, then there is no change in the determinant: $\det A' = \det A$.

  THEOREM. *Useful properties of determinants are*

  (a) $\det A^T = \det A$,

  (b) $\det(AB) = (\det A)(\det B)$,

  (c) *A is invertible if and only if* $\det A \neq 0$,

(d) *if* A *is invertible, then* $\det A^{-1} = \dfrac{1}{\det A}$ *and*

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} & \dots & C_{p1} \\ C_{12} & C_{22} & \dots & C_{p2} \\ \dots & \dots & \dots & \dots \\ C_{1p} & C_{2p} & \dots & C_{pp} \end{bmatrix}.$$

It is important to note that det is *not* a linear transformation $M_p(\mathbb{R}) \to \mathbb{R}$, as neither $\det(A+B) = \det A + \det B$ nor $\det(\alpha A) = \alpha \det A$ are true in general.

- *The Eigenvalue Problem.* If $T : V \to V$ is a linear transformation on a vector space V, then a real number $\lambda \in \mathbb{R}$ is an eigenvalue of T if there is a nonzero vector $v \in V$—called an eigenvector corresponding to λ—such that

$$Tv = \lambda v.$$

Note that $Tv = \lambda v$ if and only if

$$0 = Tv - \lambda v = Tv - \lambda I v = (T - \lambda I)v.$$

Hence, $\lambda \in \mathbb{R}$ is an eigenvalue of T if and only if $\ker(T - \lambda I) \neq \{0\}$; in such cases every nonzero vector $v \in \ker(T - \lambda I)$ is an eigenvector of T corresponding to λ. The subspace $\ker(T - \lambda I)$ is called the eigenspace corresponding to the eigenvalue λ.

If $V = \mathbb{R}^p$, then T is a $p \times p$ matrix. Therefore, it is useful to know how to find the eigenvalues of matrices $A \in M_p(\mathbb{R})$. This is achieved by way of the characteristic polynomial of A.

The characteristic polynomial of $A \in M_p(\mathbb{R})$ is the polynomial $c(\lambda)$ of degree p defined by

$$c(\lambda) = \det(A - \lambda I).$$

The roots of the characteristic polynomial of A—those $\lambda \in \mathbb{R}$ for which $c(\lambda) = 0$—are precisely the eigenvalues of A. (Why?)

Given an eigenvalue λ of A, to find the eigenvectors one solves $(A - \lambda I)x = 0$ for x. Since we are seeking the most general solutions x to $(A - \lambda I)x = 0$, we necessarily seek a basis for $\operatorname{Nul}(A - \lambda I)$.

Therefore, "to solve the eigenvalue problem $Av = \lambda v$" for $A \in M_p(\mathbb{R})$ one means:

(i) the determination of all $\lambda \in \mathbb{R}$ for which $\operatorname{Nul}(A - \lambda I) \neq \{0\}$, and

(ii) for each such λ, the determination of a basis for (the eigenspace) $\operatorname{Nul}(A - \lambda I)$.

- *Diagonalisation of Matrices.* Some (but not all!!) matrices $A \in M_p(\mathbb{R})$ have the useful property that a basis for $\mathbb{R}^p$ can found which consists entirely of eigenvectors of A. Suppose $A \in M_p(\mathbb{R})$ is such a matrix; then A is said to be diagonalisable.

If $A \in M_p(\mathbb{R})$ admits a basis $v_1, \dots, v_p$ of $\mathbb{R}^p$ in which each v_j is an eigenvector of A, then there are real numbers $\lambda_1, \dots, \lambda_p \in \mathbb{R}$ (these numbers might not be distinct—that is, there could be numbers in the list that are repeated) such that

$$Av_j = \lambda_j v_j, \quad \text{for all } j = 1, \dots, p.$$

By virtue of the fact the $v_1, \ldots, v_p$ form a basis of $\mathbb{R}^p$, the matrix $P \in M_p(\mathbb{R})$ with columns $v_1, \ldots, v_p$, namely

$$P = \begin{bmatrix} v_1 & v_2 & \cdots & v_p \end{bmatrix},$$

is invertible. Furthermore,

$$AP = \begin{bmatrix} Av_1 & Av_2 & \cdots & Av_p \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \cdots & \lambda_p v_p \end{bmatrix} = DP,$$

where $D \in M_p(\mathbb{R})$ is the diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_p \end{bmatrix}.$$

Thus,

$$P^{-1}AP = D.$$

In other words, in passing from A to $P^{-1}AP$ we have diagonalised A.

The following theorem tells us which matrices $A \in M_p(\mathbb{R})$ have the property that $\mathbb{R}^p$ has a basis of consisting of eigenvectors of A.

THEOREM. *Let $A \in M_p(\mathbb{R})$ and denote the distinct eigenvalues of A by $\lambda_1, \ldots, \lambda_q$ (and so $q \leq p$). The following statements are logically equivalent:*

(a) *A is diagonalisable;*

(b) $\dim \mathrm{Nul}(A - \lambda_1 I) + \cdots + \dim \mathrm{Nul}(A - \lambda_q I) = p$.