Notes on Linear algebra n235-1 [RLH]

Basis and dimension A vector space V has dimension n. We shall usually assume that $n < \infty$, and also that the scalars of the vector space are real numbers. If $\mathbf{u}, \mathbf{v} \in V$, then so is the linear combination $a\mathbf{u} + b\mathbf{v}$, where a and b are scalars. Every vector $\mathbf{v} \in V$ can be expressed as a linear combination of n linearly independent basis vectors $\beta = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \ldots \mathbf{e}_n\}$. We may change to another basis, but in order to represent *every* vector in V we need to have exactly n basis vectors. The vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_m\}$ are said to be 'linearly independent' if

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots a_m\mathbf{v}_m = \mathbf{0} \quad \text{implies} \quad a_1 = a_2 = \ldots a_m = 0.$$

Thus the linearly independent vectors are essentially 'different' in the sense that no one of them can be expressed as a linear combination of the others. By introducing the idea that two vectors can be orthogonal to each other (the dot product between them is zero), we can express the notion of linear independence by means of orthogonality: if a set of vectors are mutually orthogonal, then they are linearly independent. The introduction of orthogonality allows us to perform many calculations in a very nice way.

The dot product It might appear that the dot product is very special. This is true. But even in an abstract vector space (with real scalars), we can always choose a basis $\beta = \{\mathbf{e}_i\}$ in terms of which we have the representation

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ \ldots \ x_n\mathbf{e}_n.$$

This generates a correspondence between the vector $\mathbf{v}$ and its coordinate vector in R^n given explicitly by

$$[\mathbf{x}]_\beta = [\mathbf{x}] = [x_1, x_2, \ \ldots \ x_n]^t$$

A convention has crept in here: just to be definite, when we adopt matrix notation, we think of $[\mathbf{v}]$ as a $1 \times n$ matrix (a column vector). Thus, to continue, we could define $\mathbf{x} \cdot \mathbf{y}$ by $[\mathbf{x}]^t[\mathbf{y}]$ even though $\mathbf{x}$ and $\mathbf{y}$ may not themselves be vectors in R^n. We would have in this case

$$\mathbf{x} \cdot \mathbf{y} = [\mathbf{x}]^t[\mathbf{y}] = x_1y_1 + x_2y_2 + \ldots + x_ny_n = \sum_{i=1}^{n} x_iy_i.$$

The dot product If $\mathbf{x}$ is a vector in R^3, then we have $\mathbf{x} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$, and $\mathbf{y} = y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}$, where $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ is the standard basis in R^3. We define the dot product by

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + x_3y_3.$$

This definition is a bit 'dry'. We first observe that $\mathbf{x} \cdot \mathbf{x} = x_1^2 + x_2^2 + x_3^2$. By the theorem of Pythagoras we see that this is the square of the length of $\mathbf{x}$. We write the length itself using the notation $||\mathbf{x}||$ of a 'norm'. Thus in R^n we have

$$||\mathbf{x}|| = ||\mathbf{x}||_2 = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

The subscript '2' indicates that this is a special choice of a wider possible family of norms.

Matrix multiplication The dot product between two vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n can be simply written in matrix notation. First of all, we can identify $\mathbf{v}$ with the $n \times 1$ matrix $[\mathbf{v}]$ of the elements of $\mathbf{v}$. We may on occasion, if the context makes things clear, replace $[\mathbf{v}]$ with $\mathbf{v}$, or even simply with v. Anyway, it follows from the rules of matrix multiplication that $\mathbf{u} \cdot \mathbf{v} = [\mathbf{u}]^t[\mathbf{v}]$; the latter we may also write simply u^tv. By extension, we can think of the product of an $m \times k$ matrix A with a $k \times n$ matrix B as the result of $m \times n$ dot products of the m rows of A with the n columns of B. This is often a very fruitful picture.

Geometry We can introduce some more geometry into the picture by taking advantage of the **Cauchy-Schwarz inequality** which says $|\mathbf{x} \cdot \mathbf{y}| \leq ||\mathbf{x}||||\mathbf{y}||$, with equality only if $\mathbf{y}$ is a multiple of $\mathbf{x}$. If neither vector has length zero, we see that the dot product divided by the lengths is always in the range $[-1, 1]$ and we can define the angle θ between the vectors by the relation

$$\cos(\theta) = \frac{\mathbf{x} \cdot \mathbf{y}}{||\mathbf{x}||||\mathbf{y}||}.$$

This extends to vectors in R^n the notion from elementary geometry in R^3 of the angle between two vectors. Once we have this definition in place, we can define 'orthogonality' to be the case where $\theta = \pm\frac{\pi}{2}$, that is to say $\cos(\theta) = 0$, or $\mathbf{x} \cdot \mathbf{y} = 0$. This notion is only useful if neither vector has zero length.

Theorem *If the vectors* $\{\mathbf{v}_i\}$ *are mutually orthogonal, then they are linearly independent*

Proof Suppose the set of (non-zero) mutually orthogonal vectors is $\{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_m\}$, then we must consider the vanishing of a linear combination, that is to say

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots a_m\mathbf{v}_m = \mathbf{0}.$$

If we now take the dot product of the left side with the vector $\mathbf{v}_k$ we immediately obtain $a_k\mathbf{v}_k \cdot \mathbf{v}_k = a_k||\mathbf{v}_k||^2 = 0$; that is to say, $a_k = 0$. This is true for every $k, 1 \leq k \leq m$. Thus the vectors are linearly independent. $\square$

Change of basis If, in a flight of fancy, the basis vectors themselves are arranged into a column 'vector' $[e]$, then we get expressions such as $v = [e]^t[v]_e = [e]^t PP^{-1}[v]_e = [e']^t[v]_{e'}$. This little formula summarizes the whole scene to do with changes of base. The only thing left is to show that $P = [e'_i]_e$, that is to say, the columns of P are the coordinates of the new basis vectors $\{e'\}$ wrt the old ones $\{e\}$. A few bubble diagrams then lead to the usual transformation rules $[F]_{e'}^{f'} = Q^{-1}[F]_e^f P$ for transformations of the representations of linear maps. Just as with the clock changes in Spring and Autumn, this little problem can be blown quite out of proportion to its inherent difficulty. On a quiet day, one can branch here into equivalence relations and similarity transformations, thus cementing home the maxim $F \leftrightarrow [F]$, that is to say, that a linear transformation may be identified with its matrix representation with respect to bases in the domain and range: matrices are everything.

Norms A norm is a map $||\cdot||$ from vectors or matrices to the reals having 4 basic properties; in the case of square matrices a 5th (consistency) property is also included:

1. $||x|| \ \geq \ 0$
2. $||x|| \ = 0 \quad \Longleftrightarrow \quad x = 0$
3. $||cx|| \ = \ |c| \ ||x||$
4. $||x + y|| \ \leq \ ||x|| + ||y||$ (triangle inequality)

---

5. $||AB|| \ \leq \ ||A|| \ ||B||$ (for square matrices)

Inner Products If V is a vector space (over R), an inner product $\langle x, y\rangle$ is a map from $V \times V$ to the R with the following properties:

1. $\langle ax + by, z\rangle \ = \ a\langle x, z\rangle \ + \ b\langle y, z\rangle$ linearity
2. $\langle x, y\rangle \ = \ \langle y, x\rangle$ symmetry
3. $\langle x, x\rangle \ \geq \ 0$ $\langle x, x\rangle \ = \ ||x||^2$
4. $\langle x, x\rangle \ = \ 0$ if and only if $x \ = \ 0$

Cauchy-Schwarz inequality $|\langle x, y\rangle| \ \leq \ ||x|| \ ||y||.$