Mathematical Concepts used in IBIP

S.J. Maybank
7th January 2011

Department of Computer Science and Information Systems, Birkbeck College, Malet Street, London, WC1E 7HX, UK.
sjmaybank@dcs.bbk.ac.uk

This document summarizes the mathematical background of the lectures delivered in the first four weeks of the spring term for the module Image Based Information Processing.

1 Week 1

1.1 Summation

Summation is often indicated using Σ. For example,

$$\begin{aligned}\sum_{i=1}^{3} i &= 1+2+3,\\ \sum_{i=1,j=1}^{3} i \times j &= \sum_{i=1}^{3}\left(i \times \sum_{j=1}^{3} j\right) = 1\times(1+2+3)+2\times(1+2+3)+3\times(1+2+3).\end{aligned}$$

1.2 Histogram

Consider the following image.

4	5	6	1
1	4	4	6
2	6	6	5
2	2	6	6

The distinct grey levels in the image are 1, 2, 4, 5, 6. The histogram is given by the following table.

grey level	1	2	3	4	5	6
number of pixels	2	3	0	3	2	6

If the pixel values are real numbers rather than integers, then it may happen that very few of the pixel values are exactly equal. In order to construct a useful histogram it is necessary to divide the range of pixel values into a succession of intervals and then count the number of pixel values in each interval.

1.3 Linear function.

A function $f : \mathrm{I\!R}^m \to \mathrm{I\!R}^n$ is linear if for all vectors $a, b \in \mathrm{I\!R}^m$ and all real numbers r, s,

$$f(ra + sb) = rf(a) + sf(b). \tag{1}$$

Note that the property $f(0) = 0$ follows from (1) on setting $r = s = 0$. The linear function f in (1) is defined by an $n \times m$ matrix, where n is the number of rows and m is the number of columns. Conversely each $n \times m$ matrix defines a linear function from $\mathrm{I\!R}^m$ to $\mathrm{I\!R}^n$. For example, if $m = n = 2$, then the following function is linear,

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} a_1 + 2a_2 \\ a_2 \end{pmatrix}.$$

1.4 Vectors and matrices

The vector $v = (v_1, v_2)$ is a two dimensional row vector with components v_1, v_2. The corresponding column vector is $v^\top = (v_1, v_2)^\top$ where $^\top$ is transpose,

$$(v_1, v_2)^\top = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

If a vector u is introduced into a calculation, then it is usually a column vector. Let M be an $m \times n$ matrix. If the column vector u has dimension n, then M can act on u from the left to give Mu. If u has dimension m, then M can act on $u^\top$ from the right to give $u^\top M$.

The Euclidean length $\|v\|$ of the two dimensional vector $v = (v_1, v_2)$ is

$$\|v\| = (v_1^2 + v_2^2)^{1/2}.$$

The vector v can be written in the form

$$v = \|v\|(\cos(\theta), \sin(\theta)),$$

where θ is an angle that specifies the direction of v. The vector $(\cos(\theta), \sin(\theta))$ is a unit vector because

$$\|(\cos(\theta), \sin(\theta))\| = 1.$$

1.5 The Gaussian function

The definition of the Gaussian density with expected value 0 and standard deviation σ is

$$(2\pi\sigma^2)^{-1/2} \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

There is so much on the web and in text books about the Gaussian density that there seems little point in reproducing it here. Try

http://homepages.inf.ed.ac.uk/rbf/HIPR2/gsmooth.htm

which looks at Gaussian filtering and

http://en.wikipedia.org/wiki/Normal_distribution

which concentrates on the Gaussian as a probability density function.

2 Week 2

2.1 Gray level gradient

Imagine that the grey level in an image is in the form of a differentiable function

$$(x, y) \mapsto g(x, y). \tag{2}$$

Equation (2) is useful even though images are discrete rather than continuous because it gives a language for talking about gradients. The grey level gradient at (x, y) is a two dimensional vector. It points in the direction in which the grey level gradient increases most rapidly, and it has a magnitude equal to the rate of increase of grey level per unit distance in the image. The grey level gradient at (x, y) is

$$\left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right).$$

For example, if $g(x, y) = x^2 + y^2$, then

$$\left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right) = (2x, 2y).$$

2.2 Orthogonal projection

Let p be a point in the plane $\mathbb{R}^2$ and let l be a line in the plane. If p is not on l, then the orthogonal projection of p onto l is the point q on l such that the line $\langle p, q\rangle$ is perpendicular to l. If p is on l then the orthogonal projection of p onto l is p. For example, the orthogonal projection of $(0, 1)$ onto the x axis is the origin $(0, 0)$.

Let l be any line through the origin and let $v = (\cos(\theta), \sin(\theta))$ be the unit vector parallel to l. Let $p = (p_1, p_2)$ be any point in the plane and let q be the orthogonal projection of p onto l. It follows that

$$p.(\cos(\theta), \sin(\theta)) = p_1 \cos(\theta) + p_2 \sin(\theta) = \pm\|q\|.$$

(Draw a diagram!)

2.3 Scalar product of two vectors

Let a, b be vectors in n-dimensional Euclidean space $\mathbb{R}^n$ with components a_i, $1 \le i \le n$ and b_i, $1 \le i \le n$, respectively. The scalar product or dot product of a and b is defined by

$$a.b = \sum_{i=1}^{n} a_i b_i.$$

The scalar product is often written in the form $a^\top b$. The Euclidean length of a is given by

$$\|a\| = (a.a)^{1/2}.$$

It is now assumed that a and b are non-zero. Let θ be the angle between a and b. It can be shown that

$$\cos(\theta) = \frac{a.b}{\|a\|\,\|b\|}. \tag{3}$$

To prove (3), let q be the projection of a onto b. It follows that

$$\|q\| = \|b\|^{-1} a.b.$$

To complete the proof, apply to usual formula for the cosine to the triangle with vertices a, q and the origin. Note that $|\cos(\theta)| \leq 1$ and that $|\cos(\theta)| = 1$ if and only if a, b are parallel or anti-parallel.

3 Week 3

3.1 Coordinate systems

Let o be a point in $\mathbb{R}^3$ and let e_1, e_2, e_3 be an orthonormal set of vectors in $\mathbb{R}^3$, in that

$$\|e_1\| = \|e_2\| = \|e_3\| = 1,$$

and

$$e_1.e_2 = e_2.e_3 = e_3.e_1 = 0.$$

Note that $\|e_1\|^2 = e_1.e_1$, etc. Let p be a point in $\mathbb{R}^3$. The coordinates of p in the Cartesian coordinate system with origin o and coordinate axes parallel to e_1, e_2, e_3 respectively, are

$$(p - o).e_i, \qquad i = 1, 2, 3.$$

3.2 Cross product

The cross product is also known as the vector product. Let a, b be two vectors in $\mathbb{R}^3$. The cross product, $a \times b$ of a and b is a vector in $\mathbb{R}^3$. If a is 0 or b is 0, then $a \times b$ is zero. If a and b are parallel or antiparallel, then $a \times b$ is zero. In all other cases, a and b together span a unique plane in $\mathbb{R}^3$, and $a \times b$ is a certain non-zero vector perpendicular to a and b, and hence perpendicular to the plane spanned by a and b. The direction of $a \times b$ is chosen such that a, b and $a \times b$ form a right handed set of vectors in $\mathbb{R}^3$ and the magnitude of $a \times b$ is equal to

$$\|a\|\,\|b\|\,|\sin(\theta),$$

where θ is the angle between a and b.

The cross product has the following properties:

i) $a \times b = 0$ if and only if $a = b$ or $b = 0$ or a is parallel to b or a is anti-parallel to b.

ii) $(a \times b).a = 0$ and $(a \times b).b = 0$.

iii) If $a \times b \neq 0$ and if c is any vector perpendicular to a and perpendicular to b then there exists a real number λ such that $c = \lambda(a \times b)$.

iv) $a \times b = -b \times a$.

v) If a, b, c are any three vectors in $\mathbb{R}^3$, then

$$(a \times b).c = (c \times a).b = (b \times c)a.$$

3.3 Intersection of a line and a plane

A line and a plane in $\mathbb{R}^3$ intersect at a unique point unless the line is parallel to the plane. The easiest way of finding the intersection is to use a parametric version of the line. Let a be a point on the line, let u be a unit vector in the direction of the line, and let b be any other point on the line. Then there exists a unique real number $t \in \mathbb{R}$ such that

$$b = a + tu.$$

Let $p = (p_1, p_2, p_3)$ be a general point in $\mathbb{R}^3$ and let the plane have an equation

$$w.p = r,$$

where w is a non-zero vector in $\mathbb{R}^3$ and r is a real number. The point b at which the line intersects the plane is obtained by solving

$$w.(a + tu) = r \tag{4}$$

for t. If $w.u \neq 0$, then (4) has the unique solution

$$t = (r - (w.a))/w.u.$$

If $w.u = 0$, then (4) does not have a solution.

4 Week 4

4.1 Orthonormal bases

Let $e(i)$, $1 \leq i \leq n$, be vectors in $\mathbb{R}^n$ such that

$$e(i).e(i) = 1, \quad 1 \leq i \leq n,$$

and

$$e(i).e(j) = 0, \quad 1 \leq i, j \leq n \text{ and } i \neq j.$$

The vectors $e(i)$, $1 \leq i \leq n$ are said to form an orthonormal basis of $\mathbb{R}^n$. Any vector x in $\mathbb{R}^n$ can be written as a sum of the $e(i)$,

$$x = \sum_{i=1}^{n} \lambda_i e(i), \tag{5}$$

where $\lambda_i = e(i).x$ for $1 \leq i \leq n$.

Let $f(i)$, $1 \leq i \leq n$ be a second orthonormal basis for $\mathbb{R}^n$. The vector x in (5) can be written as a sum of the $f(i)$,

$$x = \sum_{i=1}^{n} \mu_i f(i).$$

Let λ be the n-dimensional vector with components λ_i, $1 \leq i \leq n$, and let μ be the n-dimensional vector with components μ_i, $1 \leq i \leq n$. Then there exists an $n \times n$ matrix M depending only on the $e(i)$ and the $f(i)$, such that

$$\mu = M\lambda. \tag{6}$$

The matrix M is orthogonal in that $MM^\top = I(n)$, where $I(n)$ is the $n \times n$ identity matrix. The orthogonal matrix M defines a rotation on $\mathbb{R}^n$.

The discrete cosine transform (DCT) is defined by an equation of the form (6). In the IBIP lectures, μ is a 1-dimensional image, i.e. a column of n pixels, and λ is the vector of DCT coefficients of μ. Where does the matrix M come from? Most textbooks obtain M from the Fourier transform but a better explanation can be found in the book "Mathematics of Digital Images" by Stuart Hoggar (CUP, 2006), see Section 15.5.