§2 Higher Dimensions.

As we remarked earlier the idea of distance between real numbers — their metrical structure — enables us to define various kinds of limits. So we can then speak of sequences as convergent, divergent and of functions being continuous, differentiable, integrable, and so on. The derivative of a function $f^{(\dagger)}$ at a point leads to the function $f'(x)$ so by taking a limit in a particular way we can associate a function with f; a function 'derived from f'. The integral of $f^{(\dagger)}$ can be arrived at by a special kind of limit also.

The idea of distance is not restricted to $\mathbb{R}$. We are familiar with it in the plane and three dimensional space, their mathematical representations are $\mathbb{R}^2$ and $\mathbb{R}^3$. So if $\underline{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ as does $\underline{y} = (y_1, y_2, y_3)$ then the distance between $\underline{x}$ and $\underline{y}$ is (using Pythagorus),

$$\|\underline{x} - \underline{y}\| = \left(\sum_{i=1}^{3} (x_i - y_i)^2\right)^{1/2}.$$

This idea extends to $\underline{x}, \underline{y}$ in $\mathbb{R}^n$:

$$\|\underline{x} - \underline{y}\| = \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{1/2}.$$

Once we have a notion of distance we can speak of:

... The sequence $(\underline{x}_n)$ converges to $\underline{x}$ iff $\forall \varepsilon > 0 \; \exists N \in \mathbb{N} : n \geq N \Rightarrow \|\underline{x} - \underline{x}_n\| < \varepsilon$ ...

The function $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous at the

(†) f is nice enough

point $\underline{x} \in \mathbb{R}^n$ iff $\forall \epsilon > 0 \ \exists \delta > 0 : \|\underline{x} - \underline{y}\|_n < \delta$ $\Rightarrow \|f(\underline{x}) - f(\underline{y})\|_m < \epsilon$. Here I've put the subscript n and m on the distance function "$\| \cdot \|$" to indicate whether it's for $\mathbb{R}^n$ or $\mathbb{R}^m$.

There are lots of different ways of measuring distance! So far we have chosen the one that we can relate to our geometric intuition. But there are other ways:

For $\underline{x}, \underline{y}$ in $\mathbb{R}^n$,

$$\|\underline{x} - \underline{y}\|_\infty = \max\{|x_1 - y_1|, |x_2 - y_2|, \dots, |x_n - y_n|\}$$

$$\|\underline{x} - \underline{y}\|_1 = \sum_{i=1}^{n} |x_i - y_i|$$

seen this already $$\|\underline{x} - \underline{y}\|_2 = \left(\sum_{i=1}^{n} |x_i - y_i|^2\right)^{1/2}$$

$$\|\underline{x} - \underline{y}\|_p = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p} \quad 1 \leq p < \infty.$$

All of these prescriptions provide a way of measuring distance in $\mathbb{R}^n$.(*) Each of them gives us a form of convergence of sequences, continuity of function and so on.

One can take these ideas further: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $f : \Omega \to \mathbb{R}$ a "random variable" ie $f^{-1}(B) \in \mathcal{F}$ for every Borel Set $B \subseteq \mathbb{R}$. We say f is in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ if

$$\int_\Omega |f|^2 \, d\mathbb{P} < \infty.$$

We define the distance between f and g in $L^2(\Omega, \mathcal{F}, \mathbb{P})$

(*) I'm asking you to believe this for now.

by

$$\|f-g\|_2 = \left(\int_\Omega |f-g|^2 d\mathbb{P}\right)^{1/2}.$$

One could define a different distance function for such f and g:

$$\|f-g\|_1 = \int_\Omega |f-g| d\mathbb{P},$$

we won't use this though; the "two-norm", $\|f-g\|_2$ will figure highly in the course.

Another example: $C[a,b]$, all continuous real valued functions on the interval $[a,b]$. An appropriate distance function is

$$\|f-g\|_\infty = \max\{|f(x)-g(x)| : x \in [a,b]\}$$

because f and g are continuous this maximum exists. With this we can speak of sequences of continuous functions converging. So (f_n) converges to f iff $\forall \epsilon > 0\ \exists N \in \mathbb{N} : n \geq N \Rightarrow \|f-f_n\|_\infty < \epsilon$. Notice that this says, "for all large enough n, f_n is uniformly close to f" — because $\max\{|f(x)-f_n(x)| : x \in [a,b]\} < \epsilon$. We can also speak of "Cauchy Sequences": (f_n) is a Cauchy Sequence iff $\forall \epsilon > 0\ \exists N \in \mathbb{N} : m,n \geq N \Rightarrow \|f_n - f_m\|_\infty < \epsilon$. Notice that this tell us immediately that for each $x \in [a,b]$, $(f_n(x))$ is a Cauchy sequence in $\mathbb{R}$. So it converges, to $f(x)$, say. Is f continuous too? Well, first of all; f takes a finite value at each $x \in [a,b]$. Moreover for $\epsilon > 0$ and $N \in \mathbb{N}$ as above, for $m,n \geq N$,

$$\|f_n - f_m\|_\infty = \max\{|f_n(x)-f_m(x)| : x \in [a,b]\} < \epsilon$$

So for each x

$$|f_n(x) - f(x)| = \lim_m |f_n(x) - f_m(x)| \leq \epsilon$$

So that $\|f_n - f\|_\infty \leq \epsilon$. This amounts to saying (f_n) converges to f in $\|\cdot\|_\infty$. Could it be that f is continuous too? Well, let $x, y \in [a,b]$. Then

$$|f(x) - f(y)| \leqslant |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)|$$

and by choosing "n" large enough we could be sure that

$$|f(x) - f(y)| \leqslant \frac{\epsilon}{3} + |f_n(x) - f_n(y)| + \frac{\epsilon}{3}$$

by having $\|f - f_n\|_\infty \leqslant \epsilon/3$ (easily done!). So now we fix just such an n. By taking $|x-y|$ small enough we know $|f_n(x) - f_n(y)| < \epsilon/3$ because f_n is a continuous function. This is enough to show f is continuous.

The space $C[a,b]$ is also separable. What this means is that there is a countable set of elements in $C[a,b]$ and any element of $C[a,b]$ is the limit(†) of a sequence selected from this countable set. There are more than one of these sort of countable sets. We call a set which is countable, and such that any $f \in C[a,b]$ is the limit of a sequence chosen from this set, a countable dense set. As we remarked, there is more than one countable dense set in $C[a,b]$. One example

(†) In $C[a,b]$, i.e. "$\|g_n - g\|_\infty \to 0$."

is given by the family of functions whose typical member is of the form that follows: Let $a = t_0 < t_1 < \dots < t_n = b$ be rational numbers(†) as are $a_0, a_1, \dots, a_n$. Set

$$\phi(t) = a_k + \left(\frac{t - t_k}{t_{k+1} - t_k}\right)(a_{k+1} - a_k) \quad \text{for } t \in [t_k, t_{k+1}]$$

There are piecewise linear functions which (might) jump at the rational points, $t_0, \dots, t_n$. Another countable dense set is furnish by the set of polynomial functions with rational coefficients: this was discovered by Weierstrass and generalised by M.H. Stone.

Metric Spaces and the Contraction Mapping Theorem.

One feature that we have ignored up to now is that all of our examples are linear spaces. Some are also algebra, that is, the have a (natural) multiplication. We turn to consider the linear structure now. Recall that a vector, or linear, space, V, is a space where you can add elements: $v, u \in V$ then $u+v \in V$ and makes good sense as an addition. There is a scalar field, for us it will be the real numbers, and the product of a vector, v, and scalar, λ, λv is an element of V. This multiplication is "nice", it obeys all the rules you would expect it to. A <u>norm</u> on a vector space V is a function

$$\| \cdot \| : V \longrightarrow [0, \infty)$$

such that, (i) $\|v\| = 0 \iff v = 0 \in V$

(ii) $\|\lambda v\| = |\lambda| \, \|v\|$ for $v \in V$ and λ a scalar

(iii) $\|u + v\| \leq \|u\| + \|v\|$ "$\Delta \leq$".

(†) I'm assuming a and b rational for convenience, not because I have to.

The function $d(u,v) = \|u-v\|$ provides us with a measure of the distance between u and v. The distance $d(u,0) = \|u\|$ is regarded as the length of u.

Ex: $\mathbb{R}^n$ with L^2 norm.

A sequence (V_n) in a Vector space with a norm, is called a Cauchy sequence iff $\forall \epsilon > 0 \ \exists N \in \mathbb{N}$: $m,n \geq N \implies \|V_n - V_m\| < \epsilon$. Just like Cauchy sequences in $\mathbb{R}$. Now in $\mathbb{R}$ every Cauchy sequence is convergent to an element of $\mathbb{R}$. We say "$\mathbb{R}$ is complete". If the same is true for our vector space; every Cauchy sequence converges to an element of V, we say that V is a complete normed vector (or linear) space. These are often called Banach Spaces. In Mathematical Finance these spaces crop up all over place. There is one type of Banach Space which is particularly prevelant. We introduce these now.

Sometimes the norm on a vector space can be obtained from another structure called an inner-product. You will have met this in $\mathbb{R}^2$: there the inner-product is called the "dot" product of vectors. We note that if $\underline{x} = (x_1, x_2)$ $\underline{y} = (y_1, y_2)$ then

$$\underline{x} \cdot \underline{y} = x_1 y_1 + x_2 y_2 .$$

This 'product' is linear in both variables, it's bilinear. Also,

$$\|\underline{x}\|_2 = (x_1^2 + x_2^2)^{1/2} = (\underline{x} \cdot \underline{x})^{1/2} .$$

So in this case the "norm is derived from the inner product". The same is true in $\mathbb{R}^n$;

$$\underline{x} \cdot \underline{y} = \sum_{i=1}^{n} x_i y_i$$

and

$$\|\underline{x}\|_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2} = (\underline{x} \cdot \underline{x})^{1/2}.$$

An in a similar fashion the norm on $L^2(\Omega, \mathcal{F}, \mathbb{P})$ comes from,

$$\langle f, g \rangle = \int_\Omega fg \, d\mathbb{P},$$

and

$$\|f\|_2 = \left(\int_\Omega |f|^2 d\mathbb{P} \right)^{1/2}.$$

The notation $\langle \cdot , \cdot \rangle$ is just a fancy way of writing the inner product of f and g.

Those Banach Spaces whose norm derives from an inner product are called Hilbert Spaces. For completeness we note that an inner product on a vector space, V, over $\mathbb{R}$, is a map

$$\langle \cdot , \cdot \rangle : V \times V \to \mathbb{R}$$

such that

(i) $\langle v, u \rangle = \langle u, v \rangle$

(ii) $\langle \lambda v + \mu w, u \rangle = \lambda \langle v, u \rangle + \mu \langle w, u \rangle$ for scalars $\lambda, \mu \in \mathbb{R}$

(iii) $\langle v, v \rangle \geq 0$ and $\langle v, v \rangle = 0$ iff $v = 0$.

[CB: NOW MENTION ORTHOGONALITY etc]

Given $h \in \mathcal{H}$, $\{h\}^\perp$ consists of all those vectors orthogonal to h; ie

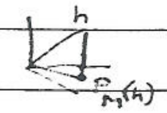
$$\{h\}^\perp = \{ k : \langle h, k \rangle = 0 \}$$

This is a <u>subspace</u> of $\mathcal{H}$ (exercise).

If $M \subseteq \mathcal{H}$ is a subspace of $\mathcal{H}$ then

$$M^{\perp} = \{k \in \mathcal{H} : \langle h, k \rangle = 0 \ \forall h \in M\}$$

This is a subspace of $\mathcal{H}$ — the subspace orthogonal to M.

[OB: 3D Preamble, nearest point  $h - P_M(h) \perp M$]

<u>Theorem</u>

Let M be a <u>closed</u> subspace of $\mathcal{H}$. For each $h \in \mathcal{H}$ there is an element of M, $P_M(h)$ say, such that

$$\|h - P_M(h)\| = \inf\{\|h - m\| : m \in M\}.$$

The element of M, $P_M(h)$, is called the orthogonal projection of h onto M.

Rem: One way of thinking of condex.

<u>Pf</u>

There is a sequence of points (m_k), in M such that $\|h - m_k\| \to \inf\{\|h - m\| : m \in M\} = \delta$ say. We show that (m_k) is a Cauchy sequence in $\mathcal{H}$. Consider $\|m_k - m_n\|$. Using the parallelogram law and observing first that;

$$\|m_k - m_n\|^2 = \|m_k - h + h - m_n\|^2$$

we get

$$\|m_k + m_n - 2h\|^2 + \|m_k - m_n\| = 2(\|m_k - h\|^2 + \|m_n - h\|^2)$$

so

$$\|m_k - m_n\|^2 = 2\left(\|m_k - h\|^2 + \|m_n - h\|^2\right) - \|m_k + m_n - 2h\|^2$$

but $\|m_k + m_n - 2h\|^2 = 4\|\frac{(m_k + m_n)}{2} - h\|^2$ which, since $\frac{m_k + m_n}{2} \in M$, is greater than $4\delta^2$. It follows that

$$\|m_k - m_n\|^2 \leqslant 2\left(\|m_k - h\|^2 + \|m_n - h\|^2\right) - 4\delta^2 \quad --- \quad (*)$$

as k and n tend to infinity $\|m_k - h\|$ and $\|m_n - h\|$ tend to δ. This shows that the right side above tends to 0. So (m_k) is Cauchy in $\mathcal{H}$. $\mathcal{H}$ is complete and therefore $m_k \to P_M(h)$ (we call it thus). Notice that since M is closed $P_M(h) \in M$. Some things follow immediately:

(i) If $h \in M$ then $P_M(h) = h$

(ii) $P_M(h)$ is unique: If m^1 and m^2 are in M and $\delta = \|h - m^1\| = \|h - m^2\|$ then using the inequality $(*)$ above, (and the same argument),

$$\|m^1 - m^2\|^2 = \|(m^1 - h) - (m^2 - h)\|^2$$

$$\leqslant 2\left(\|m^1 - h\|^2 + \|m^2 - h\|^2\right) - 4\delta^2$$

$$= 4\delta^2 - 4\delta^2 = 0.$$

So $m^1 = m^2$

(iii) $h - P_M(h)$ is orthogonal to every element of M: let $\alpha \in \mathbb{R}$, $m \in M$, then

$$\|h - P_M(h) + \alpha m\| = \|h - (P_M(h) + \alpha m)\| \geqslant d = \|h - P_M(h)\|$$

because $P_M(h) + \alpha m \in M$. So,

$$0 \leqslant \|h - P_M(h) + \alpha m\|^2 - \|h - P_M(h)\|^2 = \langle h - P_M(h) + \alpha m, h - P_M(h) + \alpha m\rangle - \langle h - P_M(h), h - P_M(h)\rangle$$

$$= \alpha^2 \|m\|^2 - 2\alpha \langle h - P_M(h), m\rangle.$$

Now this holds for every α. If $\langle h - P_M(h), m\rangle \neq 0$ then we can choose α to be of the form $\beta \langle h - P_M(h), m\rangle$ where $\beta \in \mathbb{R}$. So this last expression is

$$0 \leqslant \beta^2 |\langle h - P_M(h), m\rangle|^2 \|m\|^2 - 2\beta |\langle h - P_M(h), m\rangle|^2$$

$$= (\beta^2 \|m\|^2 - 2\beta) |\langle h - P_M(h), m\rangle|^2.$$

Taking $\beta = \frac{1}{\|m\|^2}$ yeilds $0 \leqslant \left(\frac{1}{\|m\|^2} - \frac{2}{\|m\|^2} \right) |\langle h - P_m(h), m\rangle|^2$

contradiction the right side is $-$ve. So for $m \in M$

$$0 = \langle h - P_M(h), m\rangle = \langle h, m\rangle - \langle P_M(h), m\rangle$$

ie.

$$\langle h, m\rangle = \langle P_m(h), m\rangle.$$

So $P_m(h)$ is that element of M whose inner product with every element of M agrees with that elements inner product with h. (Mention Condex Preserves Expectation)