Notes for DRE7017 Mathematics

Eivind Eriksen

Contents

① Vector spaces

<u>Defn:</u> A <u>vector space</u> is a set V, with elements $\underline{v} \in V$ called <u>vectors</u>, together with two operations

(a) <u>addition</u>: $V \times V \longrightarrow V$
$(\underline{v}, \underline{w}) \longmapsto \underline{v} + \underline{w}$

(b) <u>scalar multiplication</u>: $\mathbb{R} \times V \longrightarrow V$
$(r, \underline{v}) \longmapsto r \cdot \underline{v}$

such that the following conditions hold:

i) $(\underline{u} + \underline{v}) + \underline{w} = \underline{u} + (\underline{v} + \underline{w})$ for all $\underline{u}, \underline{v}, \underline{w} \in V$

ii) there is a <u>zero vector</u> $\underline{0} \in V$ with the property that $\underline{u} + \underline{0} = \underline{u}$ for all $\underline{u} \in V$

iii) for each vector $\underline{v} \in V$, there is an <u>additive inverse</u> $-\underline{v} \in V$ with the property that
$$\underline{v} + (-\underline{v}) = \underline{0}$$

iv) $\underline{u} + \underline{v} = \underline{v} + \underline{u}$ for all $\underline{u}, \underline{v} \in V$

v) $r \cdot (s\underline{v}) = (rs)\underline{v}$ for all $r, s \in \mathbb{R}$, $\underline{v} \in V$

vi) $(r+s) \cdot \underline{v} = r\underline{v} + s\underline{v}$ for all $r, s \in \mathbb{R}$, $\underline{v} \in V$

vii) $r \cdot (\underline{v} + \underline{w}) = r\underline{v} + r\underline{w}$ for all $r \in \mathbb{R}$, $\underline{v}, \underline{w} \in V$

viii) $1 \cdot \underline{v} = \underline{v}$ for all $\underline{v} \in V$

Example: $V = \mathbb{R}^n$

Let $V = \mathbb{R}^n$, with vectors $\underline{v} = (v_1, v_2, \dots, v_n)$ and operations defined by

i) $\underline{v} + \underline{w} = (v_1, v_2, \dots, v_n) + (w_1, w_2, \dots, w_n)$
$:= (v_1 + w_1, v_2 + w_2, \dots, v_n + w_n)$

ii) $r \cdot \underline{v} = r \cdot (v_1, v_2, \dots, v_n)$
$:= (rv_1, rv_2, \dots, rv_n)$

It is straight-forward to check that this is a vector space, with $\underline{0} = (0, 0, \dots, 0)$ and $-\underline{v} = (-v_1, -v_2, \dots, -v_n)$.

Example: $V = C(I, \mathbb{R})$ — the set of all continuous functions $f : I = [0,1] \to \mathbb{R}$

Let $V = C(I, \mathbb{R})$, where $I = [0,1]$ is the unit interval. Then V is vector space, with operations defined by

i) $(f+g)(x) = f(x) + g(x)$ for $f, g \in V$

ii) $(rf)(x) = r \cdot f(x)$ for $r \in \mathbb{R}, f \in V$

Note that if f, g are continuous fn. on I, then $f+g$ and rf are also cont. fn. on I.

Defn: A *base* of a vectorspace V is a set $B = \{\underline{v}_1, \dots, v_n\}$ of vectors in V such that

i) B is a linearly independent set of vectors

ii) $V = \text{span}(B)$.

That is, any vector $\underline{u}$ in V can be written in the form $\underline{v} = c_1 \underline{u}_1 + c_2 \underline{v}_2 + \dots + c_n \underline{u}_n$ for a unique n-tuple $(c_1, c_2, \dots, c_n) \in \mathbb{R}^n$.

Result:
Any vector space has a base. If there is a base B of V consisting on n elements, then any base of V has n elements.

Defn: The number of elements in a base of a vector space V is called the *dimension* of V, and is written $\dim(V)$.

Example:

i) $V = \mathbb{R}^n$ is a vector space of dimension n

ii) $V = C(I, \mathbb{R})$ is an infinite-dimensional vector space

iii) If $V \subseteq \mathbb{R}^n$ is a subspace such that

a) $\underline{v}, \underline{w} \in V \Rightarrow \underline{v} + \underline{w} \in V$

b) $\underline{v} \in V, r \in \mathbb{R} \Rightarrow r\underline{v} \in V$

then V is a vector space called a *linear subspace* of V, and $\dim(V) \leq n$.

<u>Defn</u>: An <u>inner product</u> on a vector space V is a product

$$V \times V \longrightarrow \mathbb{R}$$
$$(\underline{v}, \underline{w}) \longmapsto \langle \underline{v}, \underline{w} \rangle$$

that satisfies the following conditions:

i) $\langle \underline{v}, \underline{w} \rangle = \langle \underline{w}, \underline{v} \rangle$ for all $\underline{v}, \underline{w} \in V$

ii) $\langle r\underline{u} + s\underline{v}, \underline{w} \rangle = r\langle \underline{u}, \underline{w} \rangle + s\langle \underline{v}, \underline{w} \rangle$ for all $r, s \in \mathbb{R}$, $\underline{u}, \underline{v}, \underline{w} \in V$.

iii) $\langle \underline{v}, \underline{v} \rangle \geqslant 0$ for all $\underline{v} \in \underline{V}$, and $\langle \underline{v}, \underline{v} \rangle = 0$ only if $\underline{v} = \underline{0}$.

We sometimes write $\underline{v} \cdot \underline{w}$ for $\langle \underline{v}, \underline{w} \rangle$, and call it a <u>dot product</u>. A vector space with an inner product is called an <u>inner product space</u>.

<u>Example</u>: Euclidean space

Let $V = \mathbb{R}^n$. Then the product

$$\langle \underline{v}, \underline{w} \rangle = \langle (v_1, v_2, \ldots, v_n), (w_1, w_2, \ldots, w_n) \rangle$$
$$:= v_1 w_1 + v_2 w_2 + \ldots + v_n w_n$$

is an inner product, called the <u>Euclidean</u> inner product. The vector space $V = \mathbb{R}^n$ with the Euclidean inner product is called <u>Euclidean space</u>.

Example:

Let $V = \mathbb{R}^n$, and think of a vector $\underline{v} = (v_1, v_2, \ldots, v_n)$ as a column vector

$$\underline{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

Then the product defined by the matrix product

$$\langle \underline{v}, \underline{w} \rangle = \underline{v}^T A \underline{w}$$

is an inner product on $V = \mathbb{R}^n$ for any symmetric, positive definite $n \times n$-matrix A:

i) $\langle \underline{w}, \underline{v} \rangle = \underline{w}^T A \underline{v} = (\underline{w}^T A \underline{v})^T = \underline{v}^T A^T \underline{w}$
$= \underline{v}^T A \underline{w} = \langle \underline{v}, \underline{w} \rangle$ since a 1×1-matrix is its own transpose, and A is symmetric.

ii) $\langle r\underline{u} + s\underline{v}, \underline{w} \rangle = (r\underline{u} + s\underline{v})^T A \underline{w} = (r\underline{u}^T + s\underline{v}^T) A \underline{w}$
$= r\underline{u}^T A \underline{w} + s\underline{v}^T A \underline{w} = r\langle \underline{u}, \underline{w} \rangle + s\langle \underline{v}, \underline{w} \rangle$

iii) $\langle \underline{v}, \underline{v} \rangle = \underline{v}^T A \underline{v} \geq 0$, and $\underline{v}^T A \underline{v} = 0 \Rightarrow \underline{v} = \underline{0}$
holds by defn. when A is positive definite.

Notice that the special case $\underline{A = I}$ gives the Euclidean inner product:

$$\langle \underline{v}, \underline{w} \rangle = \underline{v}^T \cdot I \cdot \underline{w} = \underline{v}^T \cdot \underline{w} = (v_1\ v_2 \ldots v_n) \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$$

$$= v_1 w_1 + v_2 w_2 + \cdots + v_n \cdot w_n$$

<u>Theorem:</u> (Cauchy–Schwartz inequality)
Let V be any inner product space. Then we have

$$|\underline{v}\cdot\underline{w}| \leq \sqrt{\underline{v}\cdot\underline{v}}\cdot\sqrt{\underline{w}\cdot\underline{w}}$$

for all $\underline{v},\underline{w}\in V$.

<u>Proof:</u> Notice that $|-|$ means the absolute value of the real number $\underline{v}\cdot\underline{w}$, and that $\underline{v}\cdot\underline{v}, \underline{w}\cdot\underline{w}\geq 0$. The case $\underline{w}=\underline{0}$ is trivial, so we assume $\underline{w}\neq\underline{0}$:

Note that the projection of $\underline{v}$ onto $\underline{w}$ is given by

$$\mathrm{proj}_{\underline{w}}(\underline{v}) = \frac{\langle\underline{v},\underline{w}\rangle}{\langle\underline{w},\underline{w}\rangle}\cdot\underline{w}$$

Hence $\underline{u} = \underline{v} - \frac{\langle\underline{v},\underline{w}\rangle}{\langle\underline{w},\underline{w}\rangle}\cdot\underline{w}$ is normal to $\underline{w}$:

$$\langle\underline{u},\underline{w}\rangle = \langle\underline{v}-r\underline{w},\underline{w}\rangle = \langle\underline{v},\underline{w}\rangle - r\langle\underline{w},\underline{w}\rangle$$
$$= \langle\underline{v},\underline{w}\rangle - \langle\underline{v},\underline{w}\rangle = 0 \quad \text{with } r = \frac{\langle\underline{v},\underline{w}\rangle}{\langle\underline{w},\underline{w}\rangle}$$

such that $r\langle\underline{w},\underline{w}\rangle = \langle\underline{v},\underline{w}\rangle$. This gives

$$\langle\underline{v},\underline{v}\rangle = \langle\underline{u}+r\underline{w},\underline{u}+r\underline{w}\rangle = \langle\underline{u},\underline{u}\rangle + 2r\langle\underline{u},\underline{w}\rangle$$
$$+ r^2\langle\underline{w},\underline{w}\rangle = \langle\underline{u},\underline{u}\rangle + r^2\langle\underline{w},\underline{w}\rangle \geq r^2\cdot\langle\underline{w},\underline{w}\rangle$$
$$= \frac{\langle\underline{v},\underline{w}\rangle^2}{\langle\underline{w},\underline{w}\rangle^2}\cdot\langle\underline{w},\underline{w}\rangle = \frac{\langle\underline{v},\underline{w}\rangle^2}{\langle\underline{w},\underline{w}\rangle} \quad \text{since } \langle\underline{w},\underline{w}\rangle = 0,$$

hence $\langle\underline{v},\underline{v}\rangle\cdot\langle\underline{w},\underline{w}\rangle \geq \langle\underline{v},\underline{w}\rangle^2$. The result follows by taking square roots, since $|\langle\underline{v},\underline{w}\rangle| = \sqrt{\langle\underline{v},\underline{w}\rangle^2}$. $\square$

Remark:

For any inner product space V, the projection of a vector $\underline{v}$ on a linear subspace $W \subseteq V$ is written $proj_{\underline{v}}(w)$. It is also called the orthonormal projection.

$\underline{u} = proj_W(\underline{v})$ is the unique vector in W such that

i) $\underline{v} = \underline{u} + (\underline{v} - \underline{u})$

ii) $u \cdot (\underline{v} - \underline{u}) = 0$

(that is, $\underline{v}$ can be decomposed in a component $\underline{u}$ in W and a component $\underline{v} - \underline{u}$ normal to W).

When W is a 1-dimensional subspace of V, spanned by a vector $\underline{w}$, then we have that
$$\boxed{proj_W(\underline{v}) = \frac{\langle \underline{v}, \underline{w} \rangle}{\langle \underline{w}, \underline{w} \rangle} \cdot \underline{w}}$$
since

$\underline{u} = \dfrac{\langle \underline{v}, \underline{w} \rangle}{\langle \underline{w}, \underline{w} \rangle} \cdot \underline{w}$ means that $\underline{u} \in W$, $\underline{v} = \underline{u} + (\underline{v} - \underline{u})$

and $\underline{u} \cdot (\underline{v} - \underline{u}) = 0$ by the argument in the proof of Cauchy-Schwartz.

Defn: A norm on a vector space V is a function $\begin{aligned} V &\longrightarrow \mathbb{R} \\ v &\longmapsto \|\underline{w}\| \end{aligned}$ such that the following conditions hold:

i) $\|\underline{v}\| \geq 0$ for all $\underline{v} \in V$, and $\|\underline{v}\| = 0$ only if $\underline{v} = \underline{0}$.

ii) $\|r \cdot \underline{v}\| = |r| \cdot \|\underline{v}\|$ for all $r \in \mathbb{R}$, $\underline{v} \in V$

iii) $\|\underline{v} + \underline{w}\| \leq \|\underline{v}\| + \|\underline{w}\|$ for all $\underline{v}, \underline{w} \in V$

Example: When $V = \mathbb{R}^n$ is Euclidean space, the Euclidean norm is defined by

$$\|\underline{v}\| = \|(v_1, v_2, \ldots, v_n)\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

It is straight-forward to check that this is well-defined and satisfies the requirements for a norm.

Result: If V is an inner product space with inner product $\langle -, - \rangle$, the function

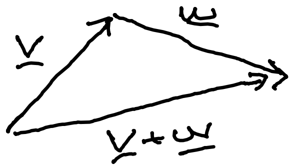
$$\|\underline{v}\| = \sqrt{\langle \underline{v}, \underline{v} \rangle}$$

is a norm on V.

We think of a norm as a way to measure the length of a vector. The third axiom of a norm

$$\|\underline{v} + \underline{w}\| \leq \|\underline{v}\| + \|\underline{w}\|$$

is called the triangle inequality. We can picture it as



when $\underline{v}, \underline{w}$ are vectors in a Euclidean space.

2. Metric spaces

Defn. A metric on a space X is a function $d: X \times X \to \mathbb{R}$ which satisfies the following:

i) $d(p,q) \geq 0$ for all $p,q \in X$, and $d(p,q) = 0$ if and only if $p = q$.

ii) $d(p,q) = d(q,p)$ for all $p,q \in X$

iii) $d(p,r) \leq d(p,q) + d(q,r)$ for all $p,q,r \in X$

Note that this definition can be applied to any space (set) X; it is not necessary that X is a vector space. We think of $d(p,q)$ as a distance between the points $p,q \in X$.

Remark:
If V is a vector space with a norm $\|\underline{v}\|$, then $d(\underline{v},\underline{w}) = \|\underline{v} - \underline{w}\|$ defines a metric on V. When $V = \mathbb{R}^n$ is Euclidean space, with the Euclidean norm, the the induced metric is called the Euclidean metric or distance, given by

$$d(\underline{v},\underline{w}) = \sqrt{(v_1-w_1)^2 + (v_2-w_2)^2 + \ldots + (v_n-w_n)^2}$$

for $\underline{v} = (v_1, \ldots, v_n)$, $\underline{w} = (w_1, \ldots, w_n)$.

Sequences

Let (X, d) be a metric space, and consider a sequence in X

$$(x_i)_{i\in\mathbb{N}} = x_1, x_2, x_3, \dots, x_n, \dots$$

with $x_i \in X$ for all $i \in \mathbb{N}$. We sometimes write (x_i) for a sequence in X.

Defn: The sequence (x_i) in X converges to a limit $x \in X$ if $d(x_i, x) \to 0$ as $i \to \infty$.
More precisely, (x_i) converges to $x \in X$ if the following condition holds:

> For any $\varepsilon > 0$, there is a natural number N such that $i > N \implies d(x_i, x) < \varepsilon$

Note that this notion of convergence depends on the metric d on X. When (x_i) converges to x, we write $(x_i) \to x$.

Ex: $X = \mathbb{R}$ with metric $d(x, y) = |x - y|$
(Euclidean space of dimension 1)

$x_i = 1/i$ for $i = 1, 2, 3, \dots$ i.e. $1, 1/2, 1/3, 1/4, \dots$

In this case, $(x_i) \to 0$ since

$$i > 1/\varepsilon \implies d(x_i, 0) = 1/i < \varepsilon$$

That is, we may choose N to be any integer greater than $1/\varepsilon$.

More informally, we say that $1/i \to 0$ when $i \to \infty$.

__Defn.__ The sequence (x_i) is __bounded__ if there is a positive number $M > 0$ in $\mathbb{R}$ and a point $p \in X$ such that $d(x_i, p) < M$ for all i. The set

$$B(p, M) = \{x \in X : d(x, p) < M\}$$

is called the __open ball__ around p with __radius__ M.

__Proposition.__ Let (x_i) be a sequence in X. Then we have:

i) If $(x_i) \to x$ and $(x_i) \to x'$, then $x = x'$.

ii) If $(x_i) \to x$, then it is bounded.

iii) If $(x_i) \to x$, then the __Cauchy criterion__ holds: For any $\varepsilon > 0$, there is an $N > 0$ such that $j, k > N \Rightarrow d(x_j, x_k) < \varepsilon$.

We say that (x_i) is a __Cauchy sequence__ if it satisfies the Cauchy criterion. It is easier to check whether a sequence is Cauchy than to check whether it converges.

Theorem:

Let (X,d) be the Euclidean space $X=\mathbb{R}^n$ with the Euclidean metric. Then every Cauchy sequence (x_i) in X converges to some element $x \in X$.

Defn. A metric space (X,d) is complete if every Cauchy sequence in X converges to a limit in X. By the theorem above, Euclidean space is complete.

Ex: Let $X = \{x \in \mathbb{R} : x > 0\} \subseteq \mathbb{R}$ with the Euclidean metric $d(x,y) = |x-y|$ for $x,y \in X$.
Then the sequence (x_i) with $x_i = 1/i$ is a Cauchy sequence in X, but it does not converge to a limit in X. In fact, $(x_i) \to 0$ in $\mathbb{R}$, but $0 \notin X$.

Defn. Let (x_i) be a sequence in X. A subsequence of (x_i) is a sequence of the form (y_j), where $y_j = x_{i_j}$ and where

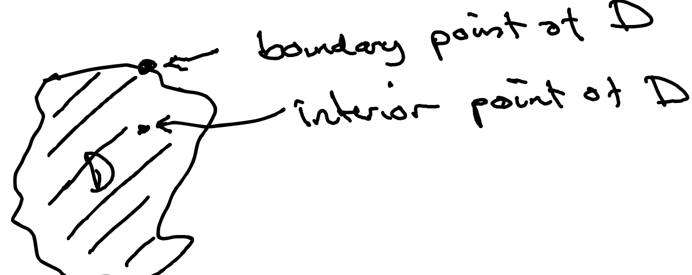
$$1 \leq i_1 < i_2 < i_3 < \cdots$$

is an infinite sequence of natural numbers.

Topology. Let (X,d) be a metric space.

Defn. A subset $U \subseteq X$ is *open* if the following condition holds: For any point $p \in U$, there is an open ball $B(p,M) \subseteq U$ for $M>0$ small enough. A subset $V \subseteq X$ is *closed* if the complement $U = X \setminus V$ is open.

Defn. Let $D \subseteq X$ be any subset of X. A *boundary point* of D is a point $p \in X$ such that any open ball $B(p,M)$ around $p \in X$, with $M>0$, contains both points in D and points in the complement $D^c = X \setminus D$. We write ∂D for the set of boundary points of D.



Defn. Let $D \subseteq X$ be any subset of X. An *interior point* of D is a point $p \in D$ that is not a boundary point of D.

Results:

① $D \subseteq X$ is closed $\Leftrightarrow$ $\partial D \subseteq D$

② $D \subseteq X$ is open $\Leftrightarrow$ all points in D are interior points

Ex: Let $X = \mathbb{R}$ with Euclidean metric. Then all open intervals (a,b) are open sets, and all closed intervals $[a,b]$ are closed sets. The boundary points of these sets are the endpoints $\{a,b\}$.

Ex: Let $X = \mathbb{R}^n$ with Euclidean metric. Then the open ball around a point $p \in \mathbb{R}^n$

$$B(p,M) = \{q \in X : d(p,q) < M\} \subseteq X$$

with radius $M > 0$ is open. The closed ball

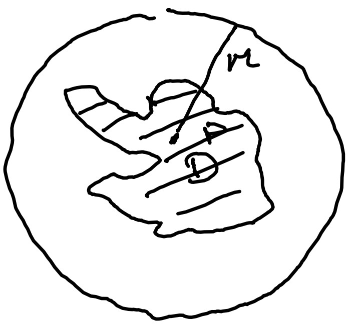
$$\overline{B}(p,M) = \{q \in X : d(p,q) \leq M\} \subseteq X$$

is closed.

Defn: Let $D \subseteq X$ be any subset. We define the closure $\overline{D} = D \cup \partial D \subseteq X$ to be the set containing D and all its boundary points. This is a closed set, and it is the minimal closed set containing D.

Ex: The closure of $D = (a,b)$ is $\overline{D} = [a,b]$ when $X = \mathbb{R}$. The closure of $D = B(p,M)$ is $\overline{D} = \overline{B}(p,M)$, the closed ball, when $X = \mathbb{R}^n$.

Defn. A subset $D \subseteq X$ is __bounded__ if there is a point $p \in D$ and a (finite) radius $M > 0$ such that $D \subseteq B(p, M)$.



D is bounded since it is contained in the open ball $B(p, M)$ for some p, $M > 0$.

Defn. A subset $D \subseteq X$ is __compact__ if the following condition holds: For any sequence (x_i) in D, there is a subsequence that converges in D.

Remarks:

i) If a sequence (x_i) in D converges to x in D, then every subsequence of (x_i) also converges to x in D.

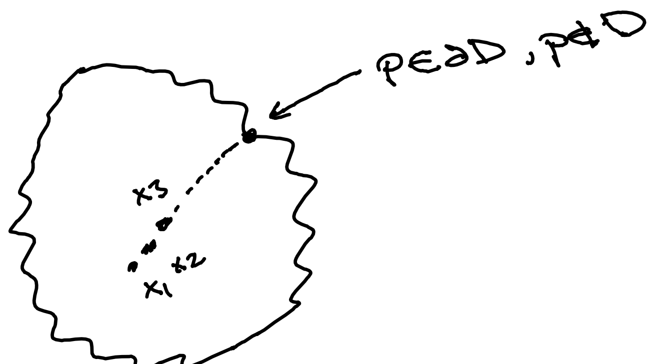
ii) The opposite implication does __not__ hold. For example, the sequence $1, -1, 1, -1, \ldots$ does not converge, but the subsequence $1, 1, 1, \ldots$ converges to 1.

Proposition. Let (X, d) be a metric space, and let $D \subseteq X$ be any subset. If D is compact, then D is closed and bounded.

Proof.

i) D compact $\Rightarrow$ D closed:

Assume D not closed. Then there is a pt. We can choose a $p \in \partial D$ such that $p \notin D$. We can choose a sequence (x_i) in D such that $(x_i) \to p$, hence D is not compact.



ii) D compact $\Rightarrow$ D bounded:

Assume D not bounded, and construct a sequence by induction such that when $x_1, \ldots, x_n \in D$ are given, then $x_{n+1} \in D$ satisfies $d(x_{n+1}, x_i) > 1$ for $i = 1, 2, \ldots, n$. This is possible since D is not bounded. Then $d(x_i, x_j) > 1$ for all i, j, hence no subsequence of (x_i) is Cauchy. This means that D is not compact; if D

was compact, then there would be a subsequence of (x_i) that was convergent, and therefore Cauchy. $\square$

<u>Theorem</u> (Bolzano-Weierstrass)

Let (X, d) be Euclidean space $X = \mathbb{R}^n$ with the Euclidean metric. Then a subset $D \subseteq X$ is compact if and only if it is closed and bounded.

3 Functions

Defn. A function $f: X \to Y$ is a rule that assigns an element $f(x)$ in Y to any element x in X. We call X the domain and Y the codomain of f.

Example:

Let $f(x,y) = e^{xy}$. We think of $f: \mathbb{R}^2 \to \mathbb{R}$ as a function with domain $\mathbb{R}^2$ and codomain $\mathbb{R}$, and it is usual to write it

$$\begin{array}{ccc} \mathbb{R}^2 & \xrightarrow{f} & \mathbb{R} \\ (x,y) & \longmapsto & e^{xy} \end{array}$$

Example: Let A be a 2×2-matrix, and let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be the function given by

$$T\begin{pmatrix} x \\ y \end{pmatrix} = A \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

By a similar construction, any $m\times n$-matrix A gives a function

$$\begin{array}{cccc} T: & \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & \underline{v} & \longmapsto & A\cdot\underline{v} \end{array}$$

Functions of this type are called linear transformations in general, and linear operators if $n=m$.

Example:

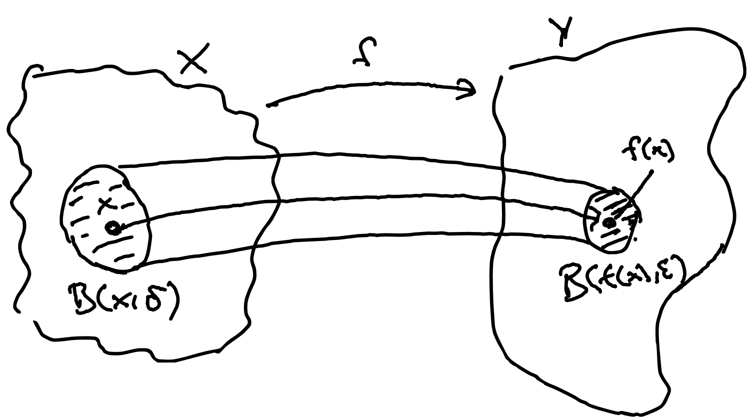
Let $\mathcal{J} : C(I, \mathbb{R}) \to \mathbb{R}$ be the function defined by $f \mapsto \int_0^1 f(x)\,dx$. This type of functions, where the domain is a set of functions, is called a functional.

Defn. Let $f: X \to Y$ be a function of metric spaces. We say that f is continuous at a point $x \in X$ if the following condition holds:

For any $\varepsilon > 0$, there exists a $\delta > 0$ such that $f(B(x, \delta)) \subseteq B(f(x), \varepsilon)$; that is, for any x' in X we have

$$d(x, x') < \delta \Longrightarrow d(f(x), f(x')) < \varepsilon$$

We say that $f: X \to Y$ is continuous if it is continuous at x for all pts $x \in X$.



Proposition:

If $f: X \to Y$ is continuous and $K \subseteq X$ is compact, then $f(K) = \{f(x) : x \in K\} \subseteq Y$ is compact.

Example:

If $f: \mathbb{R}^n \to \mathbb{R}$ is a continuous function in n variables, and $D \subseteq \mathbb{R}^n$ is closed and bounded, then it follows from the proposition that D is compact $\Rightarrow f(D)$ compact. The compact sets in $\mathbb{R}$ are closed intervals $[a,b]$ (or unions of such intervals), hence f attains a max/min on D. Hence Weierstrass' Extreme Value thm. follows from the proposition.

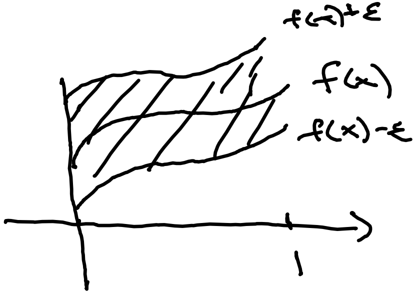
Example:

- all "usual" fn's $f: \mathbb{R}^n \to \mathbb{R}$ are continuous
- sums, differences, products, quotients, compositions, inverses of continuous fn's are continuous

Example: $C(I; \mathbb{R}) \xrightarrow{\mathcal{I}} \mathbb{R}$ is continuous when
$f \longmapsto \int_0^1 f(x)\,dx$
the metric on $C(I; \mathbb{R})$ is the metric induced by the sup-norm:

$$\|f\|_{\sup} = \sup_{x \in I} |f(x)|$$

In fact, given $f \in C(I; \mathbb{R})$, $g \in B(f, \varepsilon)$ if and only if $\sup_{x \in I} |f(x) - g(x)| < \varepsilon$, i.e. $f(x) - \varepsilon < g(x) < f(x) + \varepsilon$ for all $x \in I$. In other words, we have the following picture:



$f(x) - \varepsilon < g(x) < f(x) + \varepsilon$ means that the graph of g is in the marked area

Hence $\int_0^1 f(x) - \varepsilon \, dx < \int_0^1 g(x)\,dx < \int_0^1 f(x) + \varepsilon \, dx$,

or $J(f) - \varepsilon < J(g) < J(f) + \varepsilon$ since $\int_0^1 \varepsilon \, dx = \varepsilon$.

This implies $d(J(f), J(g)) < \varepsilon$, and therefore J is continuous (we can take $\delta = \varepsilon$).

<u>Note</u>: When $D \subseteq X$ is compact, and $C(D, \mathbb{R})$ is equipped with the sup-norm, then $C(D, \mathbb{R})$ is a <u>complete</u> metric space. That is, any Cauchy sequence in $C(D, \mathbb{R})$ converges.

Defn. Let $f: \mathbb{R}^n \to \mathbb{R}$ be a function in n variables. Then the i'th partial derivative of f is the function

$$\frac{\partial f}{\partial x_i} = f'_{x_i} = \lim_{h \to 0} \frac{f(\underline{x} + \underline{e_i} \cdot h) - f(\underline{x})}{h}$$

if the limit exists, where we write $\underline{x} = (x_1, \dots, x_n)$ and $\underline{e_i} = (0, 0, \dots, \underset{\substack{\uparrow \\ i\text{'th component}}}{1}, \dots, 0)$. The total derivative of f is a $1{\times}n$-matrix $Df(\underline{x}) = A$ such that the following condition holds:

For any $\varepsilon > 0$, there is an $\delta > 0$ such that

$$\|\underline{y} - \underline{x}\| < \delta \Rightarrow |f(\underline{y}) - f(\underline{x}) - A \cdot (\underline{y} - \underline{x})| < \varepsilon \|\underline{y} - \underline{x}\|$$

We can also write this condition as

$$\lim_{\underline{y} \to \underline{x}} \frac{\|f(\underline{y}) - f(\underline{x}) - A \cdot (\underline{y} - \underline{x})\|}{\|\underline{y} - \underline{x}\|} = 0$$

In this case, $Df(\underline{x}) = A = \left(\frac{\partial f}{\partial x_1}(\underline{x}), \dots, \frac{\partial f}{\partial x_n}(\underline{x}) \right)$, and we say that f is differentiable at $\underline{x}$.

Note:

i) If f is differentiable at $\underline{x}$, then all partial derivatives $\partial f / \partial x_i (\underline{x})$ exist.

ii) If all partial derivatives $\partial f / \partial x_i$ exists and are continuous around $\underline{x}$, then f is differentiable at $\underline{x}$

Defn. We say that f is C^1 if all partial derivatives exists and are continuous, and C^2 if also all second order partial derivatives exist and are continuous. Hence:

$$\boxed{f \text{ is } C^2 \Rightarrow f \text{ is } C^1 \Rightarrow f \text{ is continous}}$$

Thm. If $D \subseteq \mathbb{R}^n$ and $f : D \to \mathbb{R}$ is C^2, then the Hessian matrix

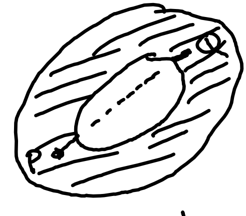
$$H(f)(\underline{x}) = \begin{pmatrix} \partial^2 f/\partial x_1^2 & \partial^2 f/\partial x_1 \partial x_2 & \cdots & \partial^2 f/\partial x_1 \partial x_n \\ \partial^2 f/\partial x_2 \partial x_1 & \partial^2 f/\partial x_2^2 & \cdots & \partial^2 f/\partial x_2 \partial x_n \\ \vdots & \vdots & \ddots & \vdots \\ \partial^2 f/\partial x_n \partial x_1 & \partial^2 f/\partial x_n \partial x_2 & \cdots & \partial^2 f/\partial x_n^2 \end{pmatrix}$$

is a symmetrie matrix.

Defn. A subset $D \subseteq \mathbb{R}^n$ is convex if for any $P, Q \in D$, the line segment $[P, Q]$ from P to Q is contained in D

D is convex

D is not convex



D is not convex

Note:

$\mathcal{D}, \mathcal{E} \subseteq \mathbb{R}^n$ convex $\Rightarrow \mathcal{D}+\mathcal{E} = \{P+Q : P \in \mathcal{D}, Q \in \mathcal{E}\} \subseteq \mathbb{R}^n$ convex

Defn: Let $f: D \to \mathbb{R}$ be a function defined on an a convex set $D \subseteq \mathbb{R}^n$. We say that:

f is convex if $\{(\underline{x}, y) \in \mathbb{R}^{n+1} : \underline{x} \in D, y \geq f(\underline{x})\}$ is a convex set

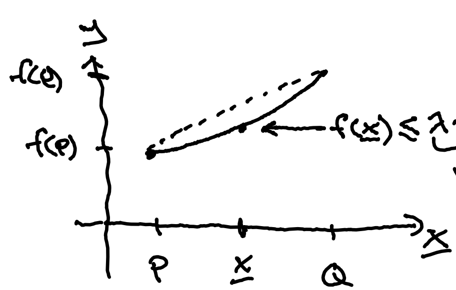
f is concave if $\{(\underline{x}, y) \in \mathbb{R}^{n+1} : \underline{x} \in D, y \leq f(\underline{x})\}$ is a convex set

Note:

i) f is concave $\Leftrightarrow -f$ convex

ii) f is convex if and only if the following condition holds:

For any pts $P, Q \in D$, consider the graph of f on $[P,Q] \subseteq D$: If $\underline{x} = \lambda \cdot P + (1-\lambda) \cdot Q \in [P,Q]$ $(0 \leq \lambda \leq 1)$, then $f(\underline{x}) \leq \lambda \cdot f(P) + (1-\lambda) \cdot f(Q)$.



f convex: $\cup$ pts "over the graph of f" form a convex set

Proposition:

If $D \subseteq \mathbb{R}^n$ is an open convex set, then a function $f : D \to \mathbb{R}$ is <u>convex</u> if and only if $H(f)(\underline{x})$ is <u>positive semidefinite</u> for all $\underline{x} \in D$.

Defn: Let $f : D \to \mathbb{R}$, where $D \subseteq \mathbb{R}^n$ is convex set.

We say that f is

<u>quasi-convex</u> if $L_f(a) = \{\underline{x} \in D : f(\underline{x}) \leq a\}$ is a convex set for all $a \in \mathbb{R}$.

<u>quasi-concave</u> if $U_f(a) = \{\underline{x} \in D : f(\underline{x}) \geq a\}$ is a convex set for all $a \in \mathbb{R}$

Example: $f(x,y) = x^2 + y^2$

f is (strictly) convex since $H(f) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ is positive definite for all (x,y) in $\mathbb{R}^2$ (and therefore positive semidefinite for all (x,y) in $\mathbb{R}^2$).

<u>quasi-convex</u>: $L_f(a) = \{(x,y) : x^2 + y^2 \leq a\} \longleftarrow z \leq a$

$\Rightarrow L_f(a)$ convex for all a

(solid circle if $a > 0$, a pt if $a = 0$, empty set if $a < 0$)

$\Rightarrow f$ is <u>quasi-convex</u>

Note: f is convex $\Rightarrow$ f is quasi-convex

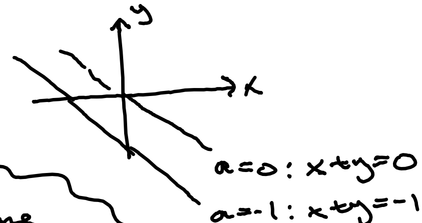
f is concave $\Rightarrow$ f is quasi-concave

Separation theorems

Defn: For any vector $\underline{p} \neq \underline{0}$ in $\mathbb{R}^n$, the set $H = \{\underline{x} \in \mathbb{R}^n : \underline{p} \cdot \underline{x} = a\}$ is called a hyperplane for any $a \in \mathbb{R}$.

Ex: $\underline{p} = (1,1)$

$H: x+y=a$

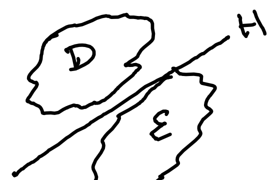


H is a line if $n=2$, a plane if $n=3$; in general a hyperplane

Defn. Two subset $D, \mathcal{E} \subseteq \mathbb{R}^n$ are separated by the hyperplane $H: \underline{p} \cdot \underline{x} = a$ if

$$D \subseteq \{\underline{x} \in \mathbb{R}^n : \underline{p} \cdot \underline{x} \geq a\}$$

$$\mathcal{E} \subseteq \{\underline{x} \in \mathbb{R}^n : \underline{p} \cdot \underline{x} \leq a\}$$



Theorem:

If $D, \mathcal{E} \subseteq \mathbb{R}^n$ are non-empty convex sets with that $D \cap \mathcal{E} = \emptyset$, then there is a hyperplane H that seperates D and $\mathcal{E}$.

Proof in the case that D is closed and $\mathcal{E} = \{\underline{x}^*\}$ is a single pt:

There is a point $\underline{y}^* \in D$ that has minimal distance to $\underline{x}^*$ among all points in D, and let

$$\underline{p} = \underline{y}^* - \underline{x}^* \neq \underline{0} \quad (\text{since } \underline{x}^* \notin D)$$

Then $H: \underline{p} \cdot \underline{x} = a$ separates D and $\mathcal{E}$, when we choose a such that $\underline{y}^* \in H$, i.e.

$$a = (\underline{y}^* - \underline{x}^*) \cdot \underline{y}^*. \qquad \square$$