Appendix A: Background Material

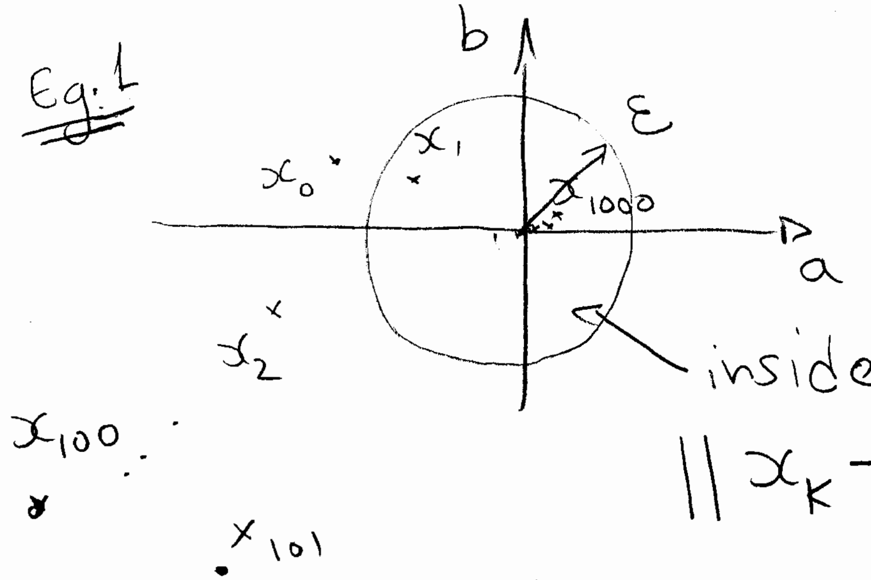
A.1 $\mathbb{R}^n$ Topology, Geometry, Analysis

Let F be a subset of $\mathbb{R}^n$, and suppose that $\{x_k\}$ is a sequence of points belonging to F.

We have: $\lim_{k \to \infty} x_k = x$, if for any $\varepsilon > 0$,

$\exists K$ s.t. $\qquad \|x_k - x\| \leq \varepsilon$, $\forall k \geq K$.

Eg. 1



inside is

$\|x_k - x\| \leq \varepsilon$,

$x = (0,0)$

For $k = 1000$, all points stay inside.

<u>Eg 2</u>

$$x_k = (1-2^{-k},\ 1/k^2)^T, \quad x = (1,0)^T$$

<u>Show</u> $x_k \longrightarrow x$

<u>proof</u>: Pick an ε. Clearly, after some K, we have; $k > K$

$$\left\| \begin{bmatrix} 1-2^{-k} \\ 1/k^2 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\|$$

$$= \left\| \begin{bmatrix} -2^{-k} \\ 1/k^2 \end{bmatrix} \right\| = \sqrt{2^{-2k} + 1/k^4} < \varepsilon$$

for "large enough" k.

$\hat{x} \in \mathbb{R}^n$ is an <u>accumulation point</u> of $\{x_k\}$ if there exist $k_1, k_2, \ldots$ such that:

$$\lim_{i \to \infty} x_{k_i} = \hat{x}.$$

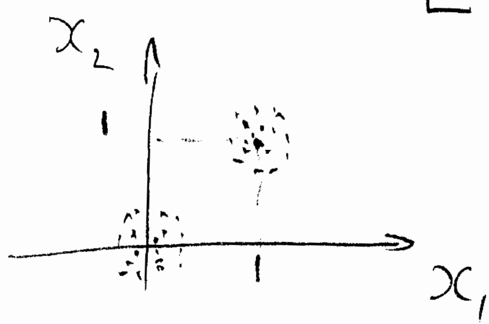
Ex 1:

$$\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1/2\\1/2\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1/4\\1/4\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1/8\\1/8\end{bmatrix}, \ldots$$

has two limit points:

$$\begin{bmatrix}0\\0\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}1\\1\end{bmatrix}$$

Picture:



Ex 2:

$x_k = \sin k$ for which, **every point** in $[-1, +1]$ is an **accumulation point**

liminf / limsup

- **liminf** gives "smallest" accumulation point.
- **limsup** gives "largest" accumulation point.

Formally:

<u>Write</u> $\hat{t} = \limsup_{k\to\infty} t_k$, $t_k \in \mathbb{R}$

<u>if</u>: (i) $\exists$ subsequence $k_1, k_2, \ldots$ (infinite) with $\lim_{i\to\infty} t_{k_i} = \hat{t}$, and

(ii) There is no other limit point t such that $t > \hat{t}$.

For liminf, replace ">" by "<".

<u>Ex</u>: $1, \frac{1}{2}, 1, \frac{1}{4}, 1, \frac{1}{8}, \ldots$

has $\liminf_{k\to\infty} t_k = 0$ and

$\limsup_{k\to\infty} t_k = 1$.

<u>Bounded Sets</u>:

A set F is bounded if $\exists M > 0$ s.t.

$$\|x\| \leq M, \quad \forall x \in F.$$

A subset $F \subset \mathbb{R}^n$ is <u>open</u> if

$\forall x \in F$, $\exists \varepsilon > 0$, <u>s.t</u>

$$\{y \in \mathbb{R}^n \mid \|x - y\| \leq \varepsilon\} \subset F$$

F is <u>closed</u> if for all possible sequences of points $\{x_k\}$ in F, <u>all limit points of</u> $\{x_k\}$ are elements of F.

<u>Exs</u>: $F = (0,1) \cup (2,0)$ is open.
$F = [0,1] \cup [2,5]$ is closed.
$F = (0,1]$ is neither open nor closed

The <u>interior</u> of a set F, denoted by int F, is the <u>largest</u> open set contained in F.

<u>Eg</u>: $\forall$ open sets $A \subset F$, $A \subseteq \text{int } F$.

The <u>closure</u> of F is the <u>smallest</u> closed set containing F.

<u>Eg</u>: ① $\forall$ closed sets A s.t. $F \subseteq A$,
$\Rightarrow \text{cl} F \subseteq A$.

② $x \in \text{cl} F$ if $\lim_{k \to \infty} x_k = x$ for some seq. $\{x_k\}$ of points in F.

Ex: $F = (-1, 1] \cup [2, 4)$

Then: $\mathrm{cl}\, F = [-1, 1] \cup [2, 4]$
$\mathrm{int}\, F = (-1, 1) \cup (2, 4)$.

Clearly: ① If F is open, then $\mathrm{int}\, F = F$.
② If F is closed, then $\mathrm{cl}\, F = F$.

Thus, for any set F, we can generate an open set $\mathrm{int}\, F$ and a closed set $\mathrm{cl}\, F$ with:

$$\mathrm{int}\, F \subseteq F \subseteq \mathrm{cl}\, F.$$

F is compact if every sequence $\{x^k\}$ of points in F has at least one limit point in F and $F \neq \emptyset$.

Very important:
$F \subset \mathbb{R}^n$ is closed and bonuded $\Longrightarrow$ F is compact.
correct the book!

* Given $x \in \mathbb{R}^n$, a neighborhood $N \subset \mathbb{R}^n$ is an <u>open set</u> containing x.

* $\mathbb{B}(x,\epsilon) = \{ y \mid \|y-x\| < \varepsilon \}$
"open Ball of radius ε around x".

* F is a <u>cone</u> if
$$x \in F \Rightarrow \alpha x \in F, \text{ all } \alpha \geq 0.$$

<u>Eg</u>: $\{ (x_1, x_2) \mid x_1 \geq 0,\ x_2 \geq 0 \}$
is a cone in $\mathbb{R}^2$

<u>correct book</u>.

* Affine hull:
$$\text{aff } F = \{ x \mid x \text{ is a linear combination of vectors in } F. \}$$

<u>Eg</u>: $F = \{ (1,0,0), (0,2,0) \}$

<u>Then</u>: $\text{aff } F = \{ (x_1, x_2, 0) \mid \forall x_1, x_2 \}$

* <u>relative interior</u> ri F is

$x \in \mathrm{ri}\, F$ if $\exists \varepsilon > 0$ s.t.

$$(x + \varepsilon B) \cap \mathrm{aff}\, F \subset F.$$

<u>Eg:1</u> $F = \{(1,0,0), (0,2,0)\}$

$\mathrm{aff}\, F = \{(x_1, x_2, 0) \mid \text{all } x_1, x_2\}$

cannot find any x so that:

$$\underbrace{(x + \varepsilon B)}_{\text{larger than points in } F.} \cap \mathrm{aff}\, F \subset F$$

$\Rightarrow \mathrm{ri}\, F = \phi$

<u>Eg 2</u>: $F = \{x \in \mathbb{R}^3 \mid x_1 \in [0,1], x_2 \in [0,1], x_3 = 0\}$

<u>Then</u> $\mathrm{aff}\, F = \mathbb{R} \times \mathbb{R} \times \{0\}$

$\mathrm{ri}\, F = \{x \in \mathbb{R}^3 \mid x_1 \in (0,1), x_2 (0,1), x_3 = 0\}$

Continuity and Limits

Let f be $f: D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$.

For some $x_0 \in \mathrm{cl}\, D$, we write

$$\lim_{x \to x_0} f(x) = f_0$$

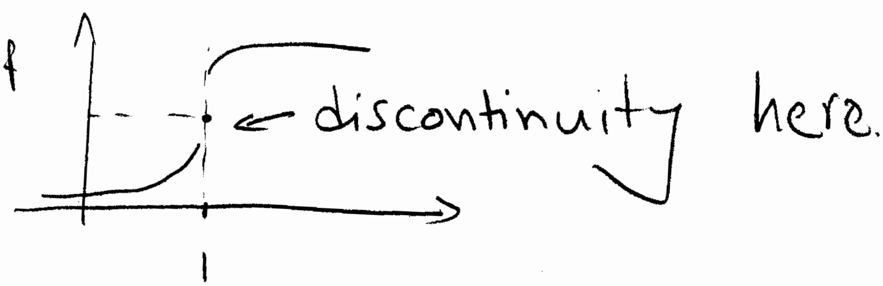
if $\forall \varepsilon > 0, \exists \delta > 0$ s.t.

$$\|x - x_0\| < \delta \text{ and } x \in D \Rightarrow \|f(x) - f_0\| < \varepsilon.$$

"You can get ε-close to f_0 by being δ-close to x_0, for some δ."

If $f(x_0) = f_0$, with $x_0 \in D$, then f is continuous at x_0.

Eg:



One-sided limits.

* For $x_0 \in \mathrm{cl}\, D$, we write:

$$\lim_{x \downarrow x_0} f(x) = f_0$$

<u>if</u> $\forall \varepsilon > 0 \ \exists \delta > 0$ s.t.

$$(x_0 < x < x_0 + \delta) \text{ and } (x \in D)$$

$$\Longrightarrow \ \|f(x) - f_0\| < \varepsilon.$$

* For $x_0 \in \mathrm{cl}\, D$, we write:

$$\lim_{x \uparrow x_0} f(x) = f_0$$

<u>if</u> $\forall \varepsilon > 0, \ \exists \delta > 0$ s.t.:

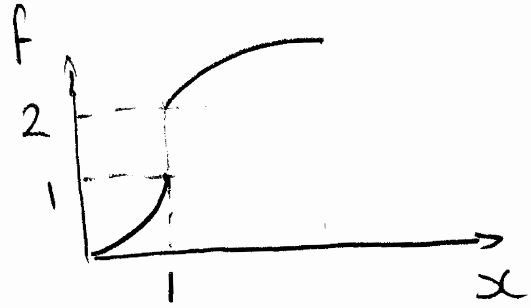
$$(x_0 - \delta < x < x_0) \text{ and } (x \in D)$$

$$\Longrightarrow \ \|f(x) - f_0\| < \varepsilon.$$

* <u>if</u> $\lim_{x \downarrow x_0} f(x) = \lim_{x \uparrow x_0} f(x) = f(x_0)$, $x_0 \in D$,

then f is conts at x_0.

Ex:



$$\lim_{x \uparrow 1} f(x) = 1, \quad \lim_{x \downarrow 1} = 2$$

$\lim_{x \to 1} f(x)$ does not exist.

* f is Lipschitz continuous if $\exists M > 0$ s.t. for any $x_0, x_1 \in D$, we have: $\|f(x_1) - f(x_0)\| \leq M \|x_1 - x_0\|$

* f is Locally Lipschitz continuous at $x_0 \in \text{int } D$ if $\exists N$, $x_0 \in N \subset D$ and $\exists M > 0$ s.t. for any $x_0, x_1 \in N$ we have: $\|f(x_1) - f(x_0)\| \leq M \|x_1 - x_0\|$

Derivatives

* $$\frac{d\phi}{d\alpha} = \phi'(\alpha) = \lim_{\varepsilon \to 0} \frac{\phi(\alpha+\varepsilon) - \phi(\alpha)}{\varepsilon}$$

* Chain rule: $$\frac{d\phi}{d\beta}(\alpha(\beta)) = \frac{df}{d\alpha} \cdot \frac{d\alpha}{d\beta}$$

* $f : \mathbb{R}^n \longrightarrow \mathbb{R}$

$$\nabla f = \begin{bmatrix} \partial f/\partial x_1 \\ \vdots \\ \partial f/\partial x_n \end{bmatrix}$$

$$\nabla^2 f = \begin{bmatrix} \partial^2 f/\partial x_1^2 & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 x_n} \\ \vdots & & & \\ \frac{\partial^2 f}{\partial x_n x_1} & & \cdots & \partial^2 f/\partial x_n^2 \end{bmatrix}$$

Note symmetry: $(\nabla^2 f)^T = \nabla^2 f$.

* For $x = x(t)$, t vector:

$$\nabla_t f(x(t)) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \nabla x_i(t)$$

Directional Derivative

$$D(f(x); p) = \lim_{\varepsilon \to 0} \frac{f(x+\varepsilon p) - f(x)}{\epsilon} = \nabla f(x)^T p.$$

Prove using formula for $\nabla_t f(x(t))$,

$\phi(\alpha) = f(y(\alpha))$, $y(\alpha) = x + \alpha p$,

and then set $\alpha = 0$ (see book).

Ex:

$$f(x_1, x_2) = x_1^2 + x_1 x_2$$
$$x_1 = \sin t_1 + t_2^2$$
$$x_2 = (t_1 + t_2)^2$$

$$\nabla_t f(x(t)) = \frac{\partial f}{\partial x_1} \cdot \nabla x_1 + \frac{\partial f}{\partial x_2} \nabla x_2$$

...

Mean value Theorem (MVT)

* Suppose that ϕ is continuous and diff'ble. Then, for $\alpha_1 > \alpha_0$, we can always find $\xi \in (\alpha_0, \alpha_1)$ s.t.:

$$\phi(\alpha_1) = \phi(\alpha_0) + \phi'(\xi)(\alpha_1 - \alpha_0)$$

* Note that the theorem restricts how much change we can have between any two points:

Let $M = \max_{\xi \in (\alpha_0, \alpha_1)} |\phi'(\xi)|$

Let $m = \min_{\xi \in (\alpha_0, \alpha_1)} |\phi'(\xi)|$

$$m(\alpha_1 - \alpha_0) < |\phi(\alpha_0) - \phi(\alpha_1)| < M(\alpha_1 - \alpha_0)$$

(Show this!)

* Multivariate extension: $f : \mathbb{R}^n \to \mathbb{R}$.

$$f(x+p) = f(x) + \nabla f(x + \alpha p)^T p$$

for some $\alpha \in (0,1)$.

MVT-2:

$$f(x+p) = f(x) + \nabla f(x)^T p + \tfrac{1}{2} p^T \nabla^2 f(x+\alpha p)^T p$$

for some $\alpha \in (0,1)$.

Implicit Function Thm

Let $h: \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^n$ satisfy:

(i) $h(z^*, 0) = 0$, some $z^* \in \mathbb{R}^n$

(ii) $h(\cdot,\cdot)$ is Lipschitz conts diff'ble in N of $(z^*, 0)$

(iii) $\nabla_z h(z,t)$ is nonsingular at $(z,t) = (z^*, 0)$

Then:

$$z: \mathbb{R}^m \rightarrow \mathbb{R}^n$$

defined implicitly by: $h(z(t), t) = 0$

is well-defined and Lipschitz conts in some neighborhood of the origin.

Geometry of Feasible Sets

→ When we look at constrained optimization.

Order Notation : $O(\cdot)$, $o(\cdot)$, $\Omega(\cdot)$

Given $\{\eta_k\}$, $\{\nu_k\}$, non-negative, we have:

① $\eta_k = O(\nu_k)$ if $\exists c$ such that: $|\eta_k| \leqslant c\,|\nu_k|$, for "sufficiently large" k.

② $\eta_k = o(\nu_k)$ if
$$\lim_{k\to\infty} \frac{\eta_k}{\nu_k} = 0$$

③ $\eta_k = \Omega(\nu_k)$ if $\exists\ 0 < c_0 \leq c_1 < \infty$ with $c_0\,|\nu_k| \leqslant |\eta_k| \leq c_1\,|\nu_k|$

(also true if $\eta_k = O(\nu_k)$ and $\nu_k = O(\eta_k)$)

Also: Ⓘ $\eta_k = O(1)$ if $|\eta_k| \leqslant c$, all k.

Ⓘⓘ $\eta_k = o(1)$ if $\lim_{k\to\infty} \eta_k = 0$

A.2 Elements of Linear Algebra

* Vectors and Matrices.
* Norms
* Subspaces
* Eigenvalues, Eigenvectors, and the SVD
* Determinant and Trace
* Matrix Factorizations: Cholesky, LU, QR.
* Error Analysis and Floating-point error
* Conditioning and Stability.