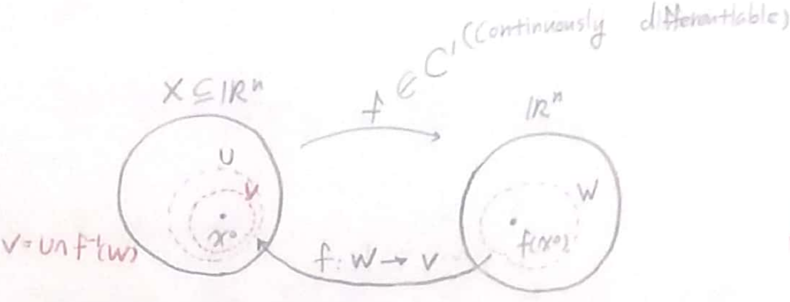
519 11/13

- Inverse function thm review



$\underline{Df^{-1}(y) = (Df(f^{-1}(y)))^{-1}}$

$Df(x^0)$ is invertible in x_0

f^{-1} locally well defined / locally C^1

$f(x) = y$: $f^{-1}(y)$ changes locally in y.

(Thm 47) < Implicit Function Theorem > $\quad \underset{\mathbb{R}^n}{X} \times \underset{\mathbb{R}^p}{\Omega} \subseteq \mathbb{R}^{n+p} \to \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix}$

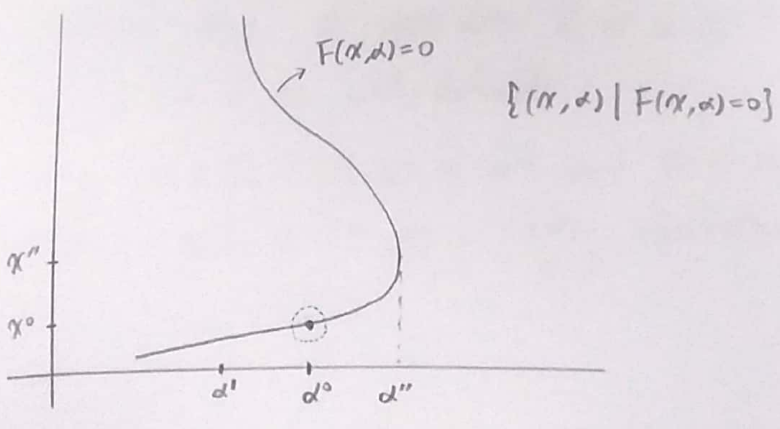
Let $X \times \Omega \subseteq \mathbb{R}^{n+p}$ be open and $F: X \times \Omega \to \mathbb{R}^n$ be continuously differentiable
Consider the system of equations $F(x, \alpha) = \underline{0}$ and assume that given some $\alpha^0 \in \Omega$
it has a solution $x^0 \in X$.

If $D_x F(x^0, \alpha^0)$ is invertible,

then (i) there exist open sets $V \subseteq \mathbb{R}^{n+p}$ and $V_\alpha \subseteq \Omega$ with $(x^0, \alpha^0) \in V$
and $\alpha^0 \in V_\alpha$ such that $\forall \alpha \in V_\alpha$, $\exists$ a unique $x_\alpha \in \mathbb{R}^n$ s.t. $(x_\alpha, \alpha) \in V$
and $F(x_\alpha, \alpha) = \underline{0}$

(ii) the solution function $x: V_\alpha \to \mathbb{R}^n$: $\alpha \longmapsto x_\alpha$ is continuously differentiable
and its derivative is given by $Dx(\alpha) = \underline{-[D_x F(x_\alpha, \alpha)]^{-1} \circ D_\alpha F(x_\alpha, \alpha)}$

(Def) $D_x F(x, \alpha) = \left[\dfrac{\partial f^i(x, \alpha)}{\partial x_j} \right]$



※ The Invese function thm as a special case of the Implicit Function theorem

If $F(x, y) := f(x) - y = 0$, with Implicit Function theorem,

$$Df^{-1}(y) = Dx(y) = [D_x(F(x_y, y))]^{-1} = (Df(f^{-1}(y)))^{-1}$$

proof of thm 4.7)

(i) Consider the function $G: X \times \Omega \to \mathbb{R}^{n+p}$ defined by

$$G^i(x,\alpha) = F^i(x,\alpha) \quad \text{for } i=1,\dots,n$$
$$G^i(x,\alpha) = \alpha_i \quad \text{for } i=n+1,\dots,n+p$$

Note that $G(x^0, \alpha^0) = \begin{pmatrix} \underline{0} \\ \alpha^0 \end{pmatrix}$ and

$$DG(x^0,\alpha^0) = \begin{pmatrix} D_x F(x^0,\alpha^0) & D_\alpha F(x^0,\alpha^0) \\ \underline{0} & I \end{pmatrix}$$

<u>Claim</u>: The associated linear operator is invertible.

Suppose $\begin{pmatrix} \lambda \\ \mu \end{pmatrix} \in \ker DG(x^0,\alpha^0)$

Then $\begin{pmatrix} D_x F(x^0,\alpha^0) & D_\alpha F(x^0,\alpha^0) \\ \underline{0} & I \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu_1 \\ \vdots \\ \mu_p \end{pmatrix} = \underline{0}$. which implies $\mu_i = 0$ for $i=1,\dots,p$

Since $D_x F(x^0,\alpha^0)$ is invertible (by assumption), this implies $\lambda_i = 0$ for $i=1,\dots,n$

Hence, we can apply the inverse function theorem.

and conclude that $\exists$ open sets $V, W \subseteq \mathbb{R}^{n+p}$ such that $(x^0,\alpha^0) \in V$, $(\underline{0},\alpha^0) \in W$ and $G^{-1}: W \to V$ is well-defined.

Let $V_\alpha = \{\alpha \in \Omega \mid (\underline{0},\alpha) \in W\}$ and $\forall \alpha \in V_\alpha$, $\exists$ a unique $x(\alpha) \in X$ such that $(x(\alpha),\alpha) \in V$ and $G^{-1}(\underline{0},\alpha) = \begin{pmatrix} x(\alpha) \\ \alpha \end{pmatrix}$ equivalent to $G(x(\alpha),\alpha) = \begin{pmatrix} \underline{0} \\ \alpha \end{pmatrix}$

(ii) The Inverse Function Theorem also implies that G^{-1} is continuously differentiable. Since $G^{-1}(\underline{0}, \alpha) = \begin{pmatrix} x(\alpha) \\ \alpha \end{pmatrix}$, this implies that $x(\alpha)$ is continuously differentiable as well.

Now, viewing G as a function of α, the chain rule and the definition of $x(\alpha)$ implies

$$DG(x(\alpha), \alpha) = \begin{pmatrix} D_x F(x(\alpha), \alpha) & D_\alpha F(x(\alpha), \alpha) \\ \underline{0} & I_p \end{pmatrix} \circ \begin{pmatrix} Dx(\alpha) \\ I_p \end{pmatrix} = \begin{pmatrix} \underline{0} \\ \underline{1} \end{pmatrix} = \begin{array}{c} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \end{array} \begin{array}{l} \}\, n \\ \\ \}\, p \end{array}$$

Hence, $D_x F(x(\alpha), \alpha) \cdot Dx(\alpha) + D_\alpha F(x(\alpha), \alpha) \cdot I = \underline{0}$

Therefore, $Dx(\alpha) = -(D_x F(x(\alpha), \alpha))^{-1} \circ D_\alpha F(x(\alpha), \alpha)$ ‖

Convex Analysis

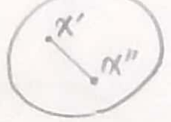
Def. • Let V be a vector space $X \subseteq V$

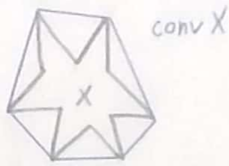
<u>X is convex</u> if, $\forall x', x'' \in X$ and $\lambda \in [0,1]$, $\lambda x' + (1-\lambda)x'' \in X$

• A point $y \in V$ is a <u>convex combination</u> of the vectors $x_1, \dots, x_m \in V$

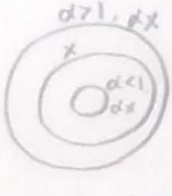
if $\exists$ $\lambda_1, \dots, \lambda_m \in [0,1]$ such that $\sum_{i=1}^{m} \lambda_i = 1$ and $y = \sum_{i=1}^{m} \lambda_i x_i$

• Let $X \subseteq V$. The <u>convex hull of X</u>, $\operatorname{conv} X$, is the smallest convex set that contains X.





ex) - If in $\mathbb{R}^n$,



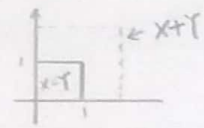
Thm 48 Let V be a vector space and $\alpha \in \mathbb{R}$

If $X, Y \subseteq V$ are convex, then $\alpha X := \{z \in V \mid \exists x \in X : z = \alpha x\}$

and $X + Y := \{z \in V \mid \exists x \in X, \exists y \in Y : z = x + y\}$ are convex

the Minkowski Sum

ex) If $X = Y$ in $\mathbb{R}^2$,



Note: We write $C - D$ as shorthand for $C + (-D)$

Thm 49 Let V be a vector space

A set $X \subseteq V$ is convex if and only if every convex combination of elements in X lies in X

Thm 50 Let V be a vector space

The convex hull of a set $X \subseteq V$ is the set of all convex combinations of elements of X.

→ Idea: Let $Y := \{y \in V : \exists x_1, \dots, x_m \in X, \lambda_1, \dots, \lambda_m \in [0,1], \sum_{i=1}^{m} \lambda_i = 1 \text{ and } \sum_{i=1}^{m} \lambda_i x_i = y\}$

One can show Y is convex which implies $Y \supseteq \operatorname{conv} X$ (Increasing case)

On the other hand, for every convex set $Z \supseteq X$, Z contains all convex combination of elements of X. (decreasing case)

Therefore, $Z \supseteq Y$, which implies $Y = \operatorname{conv} X$. ||

Thm 5.1 $X \subseteq V$ convex
$\Rightarrow$ int X and $\overline{X}$ are convex

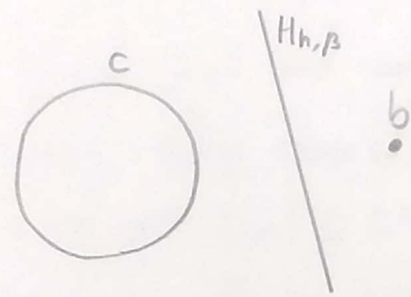
Separating Hyperplane Theorem

Def Let $h \in \mathbb{R}^n \setminus \{0\}$ and $\beta \in \mathbb{R}$.

- The set $H_{h,\beta} := \{x \in \mathbb{R}^n \mid hx = \beta\}$ is called a hyperplane
- The set $\{x \in \mathbb{R}^n \mid h \cdot x \leq \beta\}$ is called a halfspace

Thm 5.2 < Strict Separating Hyperplane theorem >

Let $C \subseteq \mathbb{R}^n$ be closed and convex and $b \in \mathbb{R}^n$, $b \notin C$.
Then, $\exists$ a hyperplane that strictly separates b and C,
that is, $\exists$ $h \in \mathbb{R}^n$ and $\overline{\beta} \in \mathbb{R}$ such that $h \cdot b < \overline{\beta}$ and $h \cdot x > \overline{\beta}$, $\forall x \in C$



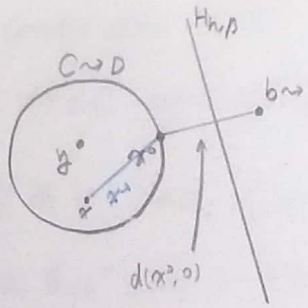
proof) We first prove the result for the translated sets $0 = b - b$ and $D = C - \{b\}$

Pick arbitrary $y \in \underset{\sim ?}{C}$ and define $D' := \{x \in D \mid d(x, 0) \leq d(y, b)\}$

Note that D' is closed and bounded and therefore compact (by Heine-Borel thm)

By the extreme value theorem, there exists $x^0 \in D'$ s.t. $d(x^0, 0) = \min_{x \in D'} d(x, 0) > 0$

Define $m = \frac{x^0}{2}$, $h = x^0$, $\beta = m \cdot h = \frac{\|x^0\|^2}{2} > 0$



Consider arbitrary $x \in D$, $x \neq x^0$ and define $\forall \lambda \in [0,1]$, $x^\lambda := x^0 + \lambda(x - x^0)$

By definition of x^0,

$$\|x^0\|^2 \leq \|x^\lambda\|^2 = (x^0 + \lambda(x - x^0))(x^0 + \lambda(x - x^0)) = \|x^0\|^2 + 2\lambda x^0(x - x^0) + \lambda^2\|x - x^0\|^2$$

This implies for all $\lambda \in (0,1)$ $0 \leq 2x^0(x - x^0) + \lambda\|x - x^0\|^2$.

Since this holds for all $\lambda > 0$, it must also hold for $\lambda = 0$.

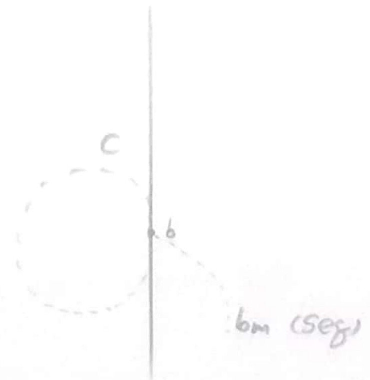
Hence, for all $x \in D$, $0 < \beta < \|x^0\|^2 \leq hx$

(Translate back)

Adding $h \cdot b$ to all inequalities and setting $\bar{\beta} := \beta + hb$,

We get $hb < \beta + hb < h(x+b)$, $\forall x \in D$

or $hb < \bar{\beta} < hy$, $\forall y \in C$. //



Thm 53 <Weak Separating Hyperplane Theorem>

Let $C \subseteq \mathbb{R}^n$ be convex and $b \in \mathbb{R}^n$, $b \notin C$

Then $\exists$ a hyperplane that separates b and C;

that is, $\exists\, h \in \mathbb{R}^n$, $\beta \in \mathbb{R}$ such that $hb \leqq \beta \leqq hx$, $\forall x \in C$

Idea) $\bar{C}$ is closed and convex, and either $b \notin \bar{C}$ or $b \in \partial C$

For $b \notin \bar{C}$ apply (thm 52) if $b \in \partial C$

$\exists$ a sequence $\{b_m\} \subseteq C^c$ that converges to b

For each b_m, thm 52 applies that $\exists\, h_m \in \mathbb{R}^n$, $\beta_m \in \mathbb{R}$ s.t. $h_m b_m < \beta_m < h_m x$, $\forall x \in C$

Assume WLOG, $\|h_m\| = 1$ Since $\{z \in \mathbb{R}^n \mid \|z\| = 1\}$ is compact.

Then, there exists a convergent subsequence h_{m_k}

$$hb = \lim_{k\to\infty} h_{m_k} \cdot b_{m_k} \leqq \lim_{k\to\infty} h_{m_k} x = hx, \ \forall x \in C,$$ setting $\beta = hb$ complete the proof

//

Thm 54 Let C and D be non-empty, disjoint and convex sets in $\mathbb{R}^n$.

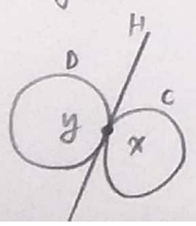
Then there is a hyperplane (h, β) such that $\forall x \in C$ and $y \in D$, $hx \geqq \beta \geqq hy$.

proof: Note that $K := C - D$ is convex and $\underline{0} \notin K$ since C and D are disjoint.

By thm 53, there exists (h, β') such that $h \cdot \underline{0} \leqq \beta' \leqq hz$, $\forall z \in K$

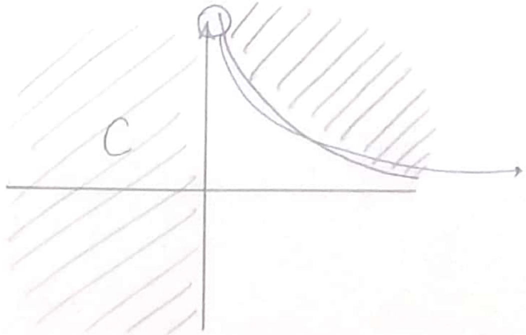
Given arbitrary $x \in C$ and $y \in D$, $z := x - y \in K$. Hence $0 \leqq hx - hy$.

This yields $hx \geqq \beta := \inf_{x' \in C} hx' \geqq hy$, $\forall x \in C$ and $y \in D$. //



Remark : Even if sets are closed, strict separation is not always possible.

ex)



finally, this part will converge so that
∄ strict hyperplanes.

(Thm55) Let C and D be two non-empty, closed, disjoint and convex sets.
Suppose C is bounded.

Then ∃ a hyperplane (h, β) such that $hx > \beta > hy$, $\forall x \in C$ and $y \in D$
(Strictly separated!)

Idea : 1) Show that $K := C - D$ is closed (Bolzano - Weierstrauss)
2) Apply strict separating hyper plane theorem.