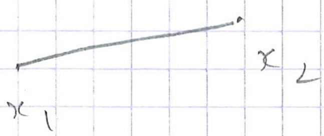
Chapter 2 / Convex Sets

§2.1 Affine & Convex Sets

§2.1.1 Line & Line Segments



Def^n: Line := For $x_1 \neq x_2 \in R^n$, $y = \theta x_1 + (1-\theta)x_2$ with $\theta \in R$ is a line.

Def^n: Line Segment := For $x_1 \neq x_2 \in R^n$, $y = \theta x_1 + (1-\theta)x_2$ with $0 < \theta < 1$ is a line segment.

Remark: $y = x_2 + \theta(x_1 - x_2)$ is also equivalent.
Start from x_2 & move along the right direction.

§2.1.2 Affine sets

Def^n: Affine set := A set C s.t. a line through any 2 points in C, lies inside C. (redundant)

:= For $x_1, x_2 \in C$, $\theta x_1 + (1-\theta)x_2 \in C$ $\forall \theta \in \mathbb{R}$.

Def^n: Affine combination := given $x_1, x_2, \cdots x_k$, the point $\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k$ s.t. $\sum \theta_i = 1$

Remark: From induction, $\Rightarrow$ an affine set has all affine combinations.

Claim: $V = C - x_0$ is a subspace, where C is an affine set, $x_0 \in C$.
Proof: Let $v_1, v_2 \in V$ & $\alpha, \beta \in R$. Now $v_1 + x_0 \in C$ & $v_2 + x_0 \in C$.
Also, $\alpha v_1 + \beta v_2 + x_0 = \alpha(v_1 + x_0) + \beta(v_2 + x_0) + (1-\alpha-\beta)x_0 \in C$
from the remark. This means $\alpha v_1 + \beta v_2 \in V$ since $(\alpha v_1 + \beta v_2) + x_0 \in C$.
[i.e. $\underbrace{\alpha v_1 + \beta v_2}_{\Downarrow} = C - x_0$]

Concl^n: Affine set C can be written as: $C = V + x_0$ (a vector space + an offset).

Def^n: Dimension of an affine set C := Dimension of the subspace $V = C - x_0$

NB: The subspace V doesn't depend on the choice of x_0

E.g.: Soln. of linear eq^{n}s is an affine set: $C = \{x \mid Ax = b\}$.

Let $x' = \theta x_1 + (1-\theta)x_2$, & x_1, x_2 be solns. Then
$Ax' = \theta Ax_1 + (1-\theta)Ax_2 = \cancel{(\theta + 1 - \theta)} b = b \Rightarrow x'$ is also a soln.
$\Rightarrow C$ is an affine set.

$\therefore$ Every affine set can be expressed as a soln. of linear eq^{n}s. (~~E~~#1)

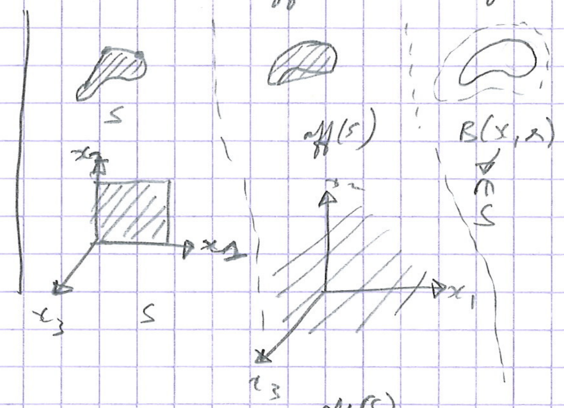
Def^n: Affine Hull := Set of all affine combinations of some set $S \subseteq \mathbb{R}^n$.
$:= aff(S) = \{\theta_1 x_1 + \dots \theta_k x_k \mid \sum \theta_i = 1,\ x_i \in S\}$
NB: $aff(S) \subseteq C$ where C is affine & $S \subseteq C$. $aff(S)$ is the smallest affine set with S in it.

§2.1.3 Affine dimension & relative interior

Def^n: Affine dim := Dimension of $aff(S)$ for any set S.
NB: Affine dim of unit circle is 2 unlike most def^{n}s.

Consider: $\dim(aff(S) \neq \mathbb{R}^n) < n$ but $S \subseteq \mathbb{R}^n$.
Def^n: Relative interior := $relint(S) = \{x \in S \mid B(x,r) \cap aff(S) \subseteq S$ for $r > 0\}$

→ Not clear, eg 2.2



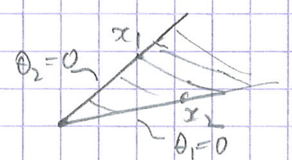
§2.1.4 Convex Sets

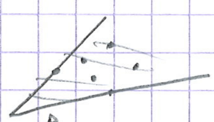
Def^n: Convex: If $\forall\, x_1, x_2 \in C$, for $0 \le \theta \le 1$, $\theta x_1 + (1-\theta) x_2 \in C$
NB: Every affine set is convex.
Def^n: Convex Combination $:= \sum \theta_i x_i$, s.t. $\sum \theta_i = 1$ & $\theta_i \geq 0\ \forall i$.

eg.

Claim: Set is convex ⇔ contains every convex combination. (If time permits)
Def^n: Convex hull := (implicit) $conv(S) = \{\sum \theta_i x_i \mid x_i \in S, \sum \theta_i = 1, \theta_i \geq 0\ \forall i\}$
Remark: Smallest convex set containing S is $conv(S)$
Generalisations: (a) Infinite sums: $\theta_i \geq 0$, $\sum_{i=1}^{\infty} \theta_i = 1$; $\sum_{i=1}^{\infty} \theta_i x_i \in S$ if converges.
(b) Integrals; $p : \mathbb{R}^n \to \mathbb{R}$,
$p(x) \geq 0\ \forall\ x \in S$, $\int_S p(x)\,dx = 1$; $\int_S p(x)\, x\, dx \in S$

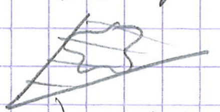
→ Not clear. (c) Probabilities:





§2.1.5 Cones

Def^n: Cone: A set S is a cone if $\forall\, x \in S$, $\theta \geq 0$ we have $\theta x \in S$.
Def^n: Convex Cone: If a set is both convex & a cone. ⇔ $\theta_1 x_1 + \theta_2 x_2 \in S$ if $\theta_i \geq 0$
Def^n: Conic Combination := (non-neg. linear combination) $\theta_1 x_1 + \dots + \theta_k x_k$ for $\theta_i \geq 0$.
Def^n: Conic Hull := (implicit)



§2.2 Some Important Examples

* Affine ($\Rightarrow$ convex): (a) ϕ (null set)
(b) $\{x_0\}$ (singleton)
(c) $\mathbb{R}^n$

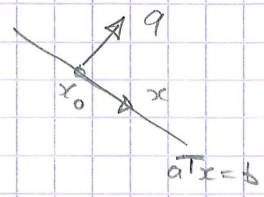
* (a) line — affine
(b) line through the origin — subspace & also convex

* line segment — convex (not affine unless a point)

* ray, $\{x_0 + \theta v \mid \theta \geq 0\}$, $v \neq 0$; — convex (not affine)
convex cone if $x_0 = 0$.

* subspace — affine, convex & a convex cone.

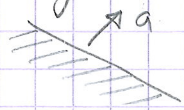
§2.2.1 Hyperplanes & halfspaces



Def^n: Hyperplane $:= \{x \mid a^T x = b\}$ where $a \in \mathbb{R}^n$, $a \neq 0$ & $b \in \mathbb{R}$

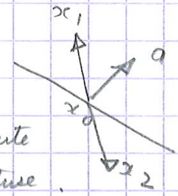
Remarks: It is a soln. of a set of linear eq^{n}s. Thus its convex.
: Recall the geometric interpretation
: can write as $\{x \mid a^T(x - x_0) = 0\}$ $(a^T x_0 = b)$
: also as $x = x_0 + a_\perp$ where $a_\perp = \{v \mid a^T v = 0\}$ to satisfy the same eq^n.

: The hyperplane is an offset x_0 plus all vectors orthogonal a (normal vector).



Def^n: Halfspaces (closed) $:= \{x \mid a^T x \leq b\}$

Remarks: Halfspaces are convex but not affine.
: The $\{x \mid a^T(x - x_0) \leq 0\}$ rep helps say that the angle of $x - x_0$ must be obtuse (or right).



a & $x_1 - x_0$ acute
a & $x_2 - x_0$ obtuse.

§2.2.2 Eucledian ball & ellipsoids

Def^n: (Eucledian) ball in $\mathbb{R}^n := B(x_c, r) = \{x \mid \|x - x_c\|_2 \leq r\} = \{x \mid (x - x_c)^T(x - x_c) \leq r^2\}$
for $r > 0$, $\|\cdot\|_2$ is the eucledian norm ($\|u\|_2 = \sqrt{u^T u}$)

Remarks: x_c is the centre, r is the radius.
: another rep $B(x_c, r) = \{x_c + ru \mid \|u\|_2 \leq 1\}$

Claim: A eucledian ball is a convex set.

Proof: Let $\|x_1 - x_c\| \leq r$, $\|x_2 - x_c\| \leq r$. For $0 \leq \theta \leq 1$,
$$\begin{aligned}\|\theta x_1 + (1-\theta)x_2 - x_c\| &= \|\theta(x_1 - x_c) + (1-\theta)(x_2 - x_c)\| \\ &\leq \theta\|x_1 - x_c\| + (1-\theta)\|x_2 - x_c\| \\ &\leq r\end{aligned}$$

Def^n: Ellipsoids $:= \mathcal{E} = \{x \mid (x - x_c)^T P^{-1}(x - x_c) \leq 1\}$ for $P = P^T \succ 0$, i.e. P is symmetric & positive definite.

Claim: Semi-axis lengths of $\mathcal{E}$ are $\sqrt{\lambda_i}$ where λ_i are eigenvalues of P.

Proof: <hint> $P = r^2 I$ yields a ball; explain also the P^{-1} instead of P.

Remark: <missing>

§2.2.3 Norm balls & norm cones.

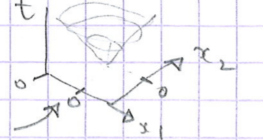
† Def^n Norm := $f: \mathbb{R}^n \to \mathbb{R}$ is a norm if
* $f \geq 0 \ \forall x \in \mathbb{R}^n$
* $f(x) = 0$ only if $x = 0$; $x = 0 \Rightarrow f(x) = 0$
* $f(tx) = |t| f(x)$
* $f(x + y) \leq f(x) + f(y)$, $\forall x, y \in \mathbb{R}$

Def^n: Norm Ball := Radius r & centre x_c $\{x \mid \|x - x_c\| \leq r\}$

Claim: Norm balls are convex

Def^n: Norm Cone $:= \{(x, t) \mid \|x\| \leq t\} \subseteq \mathbb{R}^{n+1}$

e.g. Boundary of second-order cone in $\mathbb{R}^3$ $\{(x_1 x_2 t) \mid (x_1^2 + x_2^2)^{1/2} \leq t\}$

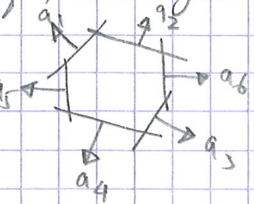


§ 2.2.4 Polyhedra

Def^n: polyhedron := $P = \{x \mid a_j^T x \leq b_j,\ j = 1 \dots m;\ c_j^T x = d_j,\ j = 1 \dots p\}$
Remarks: It's an intersection of halfspaces & hyperplanes.
e.g. affine sets (eg. subspaces, hyperplanes, lines), rays, line segments & halfspaces
Def^n: polytope := a bounded polyhedron.
Not^n: polyhedron := $P = \{x \mid Ax \preceq b,\ Cx = d\}$ with

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix}, \quad C = \begin{bmatrix} c_1^T \\ \vdots \\ c_p^T \end{bmatrix}$$

& the symbol $\preceq$ denotes component wise inequality in $\mathbb{R}^m$;



Simplexes

Def^n: Affinely independent := $v_0, \dots v_k$ are so if $v_1 - v_0, \dots v_k - v_0$ are linearly independent.

Def^n: Simplex := for $v_0, \dots v_k \in \mathbb{R}^n$ affinely independent, $S = \mathrm{conv}\{v_0 \dots v_k\}$
$= \{\theta_0 v_0 + \dots \theta_k v_k \mid \theta \succeq 0,\ \mathbb{1}^T\theta = 1\}$
$\mathbb{1} = [1, 1, 1, \dots 1]$

Remark: Affine dimension is k for the aforesaid.
E.g. 1-dimensional simplex: line segment. 3-dimensional symplex: tetrahedron
2-dimensional simplex: Triangle (plus interior)

Def^n: Unit simplex := n-dimensional simplex determined by $0, e_1 \dots e_n \in \mathbb{R}^n$
$x \succeq 0,\ \mathbb{1}^T x = 1$ ↳ vector.

Def^n: Probability simplex := <skipped>

Aim: To describe a simplex as a polyhedron; Proof: (a) $x \in C$ iff $x = \theta_0 v_0 + \dots + \theta_k v_k$ for $\theta \succeq 0$ & $\mathbb{1}^T\theta = 1$
(b) equivalently, with $y = (\theta_1, \dots \theta_k)$, $B = [v_1 - v_0, \dots v_k - v_0] \in \mathbb{R}^{n \times k}$, we have $x \in C$ iff $x = v_0 + By$ (#4) for $y \succeq 0$ with $\mathbb{1}^T y \leq 1$
(c) B has rank k ∵ $v_0, v_1 \dots v_k$ are affinely independent. ⇒ ∃ a non-singular matrix, s.t. $AB = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} B = \begin{bmatrix} I \\ 0 \end{bmatrix}$ (Row echelon form?)
(d) $A\{x = v_0 + By\}$ yields
$A_1 x = A_1 v_0 + y$, $A_2 x = A_2 v_0$ ⇒ $x \in C$ iff (i) $y = A_1 x - A_1 v_0$ granted $y \succeq 0$, $\mathbb{1}^T y \leq 1$
(ii) $A_2 x = A_2 v_0$

These are just a set of linear equalities & inequalities in x. Thus this is a polyhedron.

Convex Hull Description of Polyhedra

Recall: $\mathrm{Conv}\{v_1 \dots v_k\} = \{\theta_1 v_1 + \dots \theta_k v_k \mid \theta \succeq 0,\ \mathbb{1}^T\theta = 1\}$

ASK
RETRY

§ 2.2.5 The positive semi-definite cone.

1. Def^n: $S^n := \{X \in \mathbb{R}^{n\times n} \mid X = X^T\}$ (symmetric matrices)
Remark: S^n has dimension $n(n+1)/2$ ⚠ #3
1. Def^n: $S^n_+ := \{X \in S^n \mid X \succeq 0\}$ (symmetric positive semi definite matrix)
1. Def^n: $S^n_{++} := \{X \in S^n \mid X \succ 0\}$ (symmetric positive definite matrix)
Remark: Analogous to R_+ & R_{++}
Recall: Positive Semi definite (PSD) := if $\forall x \in \mathbb{R}^n$ $x^T A x \geq 0$, then A is PSD.
Claim: S^n_+ is a convex cone: if $\theta_1, \theta_2 > 0$, & $A, B \in S^n_+$, then $\theta_1 A + \theta_2 B \in S^n_+$.
Proof: $x^T(\theta_1 A + \theta_2 B)x = \theta_1 x^T A x + \theta_2 x^T B x \geq 0$ of course, $A \succeq 0$ & $B \succeq 0$ with $\theta_1, \theta_2 \geq 0$.

E.g. $X = \begin{bmatrix} x & y \\ y & z \end{bmatrix} \in S^2_+ \Leftrightarrow x \geq 0,\ z \geq 0,\ xz \geq y^2$ ⚠#4

§ 2.3 Operations that preserve convexity

Motivat^n: These & simple examples as described should help determine / establish convexity of sets.

§ 2.3.1 Intersection

Oct 11, 2016

Claim: Convexity is preserved by intersections; if S_1 & S_2 are convex $\Rightarrow S_1 \cap S_2$ is also convex
Remark: Extends to infinite sets; if S_α is convex, $\bigcap_\alpha S_\alpha$, $(\alpha \in A)$ is also convex.
E.g.: (1) A polyhedron – its convex. It is made of intersection of halfspaces & hyperplanes. both are convex.
(2) Positive Semi Definite Cones (S^n_+) – The can be written as $\bigcap_{z\neq 0} \{X \in S^n \mid z^T X z \geq 0\}$
– The set $\{X \in S^n \mid z^T X z \geq 0\}$ is a halfspace (since the constraint is linear) in S^n. It is hence convex.
– Concl^n: The positive semi definite cone is, itself convex b/s an intersection of convex sets.
(3) Consider set $S = \{x \in \mathbb{R}^m \mid |p(t)| \leq 1$ for $|t| \leq \pi/3\}$ for $p(t) = \sum_{k=1}^m x_k \cos kt$
<useless example, skipped writing it :)>
Claim: Every closed convex set is an intersection of halfspaces; $C = \bigcap\{\mathcal{H} \mid \mathcal{H}$ halfspaces $C \subseteq \mathcal{H}\}$
Doubt: Proof? – Mathew gave one: $x, y \in A \cap B \Rightarrow x \in A, B$ & $y \in A, B$. Now for $0 \leq \theta \leq 1$, $\theta x + (1-\theta) y \in A, B$ from convexity of A & B. obviously then $\theta x + (1-\theta)y \in A \cap B$. $\Rightarrow A \cap B$ is convex.

§ 2.3.2 Affine Functions

Def^n: Affine function := $f: \mathbb{R}^n \to \mathbb{R}^m$ is affine if it has the form $f(x) = Ax + b$ where $A \in \mathbb{R}^{m\times n}$ & $b \in \mathbb{R}^m$.
Claim: Let $S \subseteq \mathbb{R}^n$ be a convex set & let f be affine. Then $f(S)$ is also convex.
∵ if f is affine, the inverse image of S is convex; viz. $f^{-1}(S) = \{x \mid f(x) \in S\}$ is also convex.
Proof: Let $x_1, x_2 \in S \Rightarrow \theta x_1 + (1-\theta) x_2 \in S$ $(0 \leq \theta \leq 1)$.
$f(x_1), f(x_2) \in f(S)$. Also $\underbrace{f(\theta x_1 + (1-\theta) x_2)}_{\in f(S)} = \theta A x_1 + \theta c + (1-\theta) A x_2 + (1-\theta)c$
$= \theta \underbrace{f(x_1)}_{\in f(S)} + (1-\theta) \underbrace{f(x_2)}_{\in f(S)}$ □

∴ ⚠ (i think $f^{-1}(x) = \cancel{A^{-1}}^{B} x - \cancel{A^{-1}b}^{C}$ & then the usual proof can be used)
e.g. (1) Scaling & Translation: αS & $S + a$
(2) Projection: If $S \subseteq \mathbb{R}^m \times \mathbb{R}^n$, then $\{x_1 \in \mathbb{R}^m \mid (x_1, x_2) \in S$ for some $x_2 \in \mathbb{R}^n\}$

Claim: The sum of 2 convex sets is also convex, $S_1 + S_2 = \{x + y \mid x \in S_1, y \in S_2\}$ is convex if S_1, S_2 are.
Proof: Claim: If S_1 & S_2 are convex, then so is their Cartesian product.
Proof: For $(x_1, x_2) \in S_1 \times S_2$ & $(y_1, y_2) \in S_1 \times S_2$, $(\theta x_1 + (1-\theta) y_1, \theta x_2 + (1-\theta) y_2) \in S_1 \times S_2$
∵ convexity of S_1 & S_2.

Now a linear f^n on $S_1 \times S_2$, $f(x_1, x_2) = x_1 + x_2$ (is also affine). Thus, the result follows from the result (claim 1 in §2.3.2).

$\square$

Def^n: PartialSum := For $S_1, S_2 \in \mathbb{R}^n \times \mathbb{R}^m$, $S = \{(x, y_1 + y_2) \mid (x, y_1) \in S_1, (x, y_2) \in S_2\}$
Claim: Partial sums of convex sets are convex. (proof)

e.g. (3) Polyhedron; recall: $\{x \mid Ax \preceq b, Cx = d\}$; this can be expressed as an inverse image of an affine f^n, with a domain in $\mathbb{R}_+^n$ & $\{0\}$, as follows.
$\{x \mid Ax \preceq b, Cx = d\} = \{x \mid f(x) \in \mathbb{R}_+^n \times \{0\}\}$ where $f(x) = (b - Ax, d - Cx)$

(4) Sol^n set of a linear matrix inequality: The soln. set of LMI is convex where LMI is
Def^n: Linear Matrix inequality: $A(x) = x_1 A_1 + \ldots + x_n A_n \preceq B$, $A_i, B \in S^m$.
(linear in x)
Proof: Use $f(x) = B - A(x)$ (an affine f^n) & the positive semidefinite cone & $f: \mathbb{R}^n \to S^m$ then it will easily follow.

(5) Hyperbolic cone (Def^n :=) $\{x \mid x^T P x \le (c^T x)^2, c^T x \geq 0\}$ where $P \in S_+^n$ & $c \in \mathbb{R}^n$
Claim: Hyperbolic cone is convex.
Proof: Recall that the second order cone is convex, i.e. $\{(z,t) \mid z^T z \le t^2, t \geq 0\}$
Use $f(x) = (P^{1/2} x, c^T x)$ & then its obvious ($z = P^{1/2} x$, $t = c^T x$)

§2.3.3 Linear-fractional & Perspective functions

Def^n: Perspective f^n := $P: \mathbb{R}^{n+1} \to \mathbb{R}^n$ with domain $\mathbb{R}^n \times \mathbb{R}_{++}$ & $P(z,t) := \frac{z}{t}$
Remark: The perspective function scales everything to the last component & drops it.
Claim: If C ($\subseteq$ dom P) is convex, then $P(C)$ is also convex.
: Perspective f^ns preserve convexity.
Proof: Let $x = (\tilde{x}, x_{n+1})$, $y = (\tilde{y}, y_{n+1}) \in \mathbb{R}^{n+1}$ with $x_{n+1}, y_{n+1} > 0$. For $0 < \theta < 1$, we have
$$P(\theta x + (1-\theta) y) = \frac{\theta \tilde{x} + (1-\theta)\tilde{y}}{\theta x_{n+1} + (1-\theta) y_{n+1}} = \mu P(x) + (1-\mu) P(y) \text{ where } \mu = \frac{\theta x_{n+1}}{\theta x_{n+1} + (1-\theta) y_{n+1}}$$
(not too hard to see)
Claim: θ & μ are monotonically related. , for $\theta = 0, \mu = 0$; $\theta = 1, \mu = 1$.
So then convexity follows from def^n of it if x & y are chosen from a convex set.
Just use $1 = \frac{\theta x_{n+1}}{\theta x_{n+1}} > \frac{\theta x_{n+1}}{\theta x_{n+1} + (1-\theta) y_{n+1}}$ & that the denominator only inc. or dec. with θ compared to the rest.

Claim: The inverse image P^{-1} is also convex. i.e.
for $C \subseteq \mathbb{R}^n$ convex, $P^{-1}(C) = \{(x,t) \in \mathbb{R}^{n+1} \mid x/t \in C, t > 0\}$ is convex.
Proof: Suppose: $(x,t), (y,s) \in P^{-1}(C)$ & $0 \le \theta < 1$. We must show that
$\theta(x,t) + (1-\theta)(y,s) \in P^{-1}(C)$, viz. $\frac{\theta x + (1-\theta) y}{\theta t + (1-\theta) s} \in C$ ($\theta t + (1-\theta)s > 0$ is obvious)
Recall $= \mu\left(\frac{x}{t}\right) + (1-\mu)\left(\frac{y}{s}\right)$ where $\mu = \frac{\theta t}{\theta t + (1-\theta) s}$

Linear-fractional functions

Def^n: Linear-fractional function := composition of perspective f^ns with an affine f^n.
: let $g: \mathbb{R}^n \to \mathbb{R}^{m+1}$ be affine, i.e. $g(x) = \begin{bmatrix} A \\ c^T \end{bmatrix} x + \begin{bmatrix} b \\ d \end{bmatrix}$ where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ & $d \in \mathbb{R}$
$f: \mathbb{R}^n \to \mathbb{R}^m$ given by $f = P \circ g$, i.e. $f(x) = \frac{(Ax + b)}{(c^T x + d)}$, dom$(f) = \{x \mid c^T x + d > 0\}$

Remark: One can write the linear-fractional f^n as $Q = \begin{bmatrix} A & b \\ c^T & d \end{bmatrix} \in \mathbb{R}^{(m+1) \times (n+1)}$
so that $(x, 1)$ yields $(Ax + b, c^T x + d)$. Now the last component is normalized to one.
: One can think of $\mathbb{R}^n$ associated with rays in $\mathbb{R}^{n+1}$ as $P(z) = \{t(z,1) \mid t > 0\}$ for $z \in \mathbb{R}^n$ & $t \in \mathbb{R}$. Conversely any ray in $\mathbb{R}^{n+1}$ through the origin can be written as $P(v) = \{t(v,1) \mid t > 0\}$. P & $\mathbb{R}^n$ are one-one & onto. (more details).

Claim: The linear-fractional f^n can be expressed as $f(x) = P^{-1}(Q\,P(x))$ with Q & P as defined earlier.
Proof: $P(x)$ for $x \in \mathrm{dom}\, f$ viz. with $c^T x + d > 0$ will yield a ray $\{t(x,1) \mid t > 0\}$ which is then acted upon by Q to yield $t(Ax+b), t(c^Tx+d)$ which upon inversion, P^{-1}, yields the correctly transformed x. So its trivial really.

Claim: The linear-fractional f^n preserves convexity; so does its inverse.
Proof: Affine map is convex, projection map is convex.

e.g. conditional probabilities: Let $u \in \{1 \ldots n\}$, $v \in \{1, \ldots m\}$ be random variables & $p_{ij} = \mathrm{Prob}(u=i, v=j)$. $f_{ij} = Pr(u=i \mid v=j) = \frac{p_{ij}}{\sum_k p_{kj}}$

is a convex set, if the joint is a convex set.

§2.4 Generalized Inequalities

§2.4.1 Proper cones & generalized inequalities

Def^n: Proper cone := a cone $K \subseteq \mathbb{R}^n$ is proper if
- K is convex
- K is closed
- K is solid (non-empty interior)
- K is pointed (contains no line; $x \in K, -x \in K \Rightarrow x = 0$)

Motivat^n: Use this to define generalized inequalities.

18 Oct 2016

Def^n: Generalized (non-strict) inequality $:= x \preceq_K y \Leftrightarrow y - x \in K$
$:= y \succeq_K x \Leftrightarrow x \preceq_K y$

Def^n: Generalized (strict) inequality $:= x \prec_K y \Leftrightarrow y - x \in \mathrm{int}\, K$

Recall: $\mathbb{R}_+$ is convex, (closed? solid? 8) & contains no line (pointed) $\Rightarrow \mathbb{R}_+$ is a proper cone.
E.g.(1): $K = \mathbb{R}_+$ yields $\preceq_K$ as $\leq$ the usual ordering on reals.
E.g.(2): S^n_+ is a proper cone $\subset S^n$. $X \prec_K Y \Rightarrow Y - X \in S^n_+$. K is often dropped.
E.g.(3): Claim: Cone of "polynomials non-negative on $[0,1]$" is proper: $K = \{c \in \mathbb{R}^n \mid c_1 + c_2 t + \ldots c_n t^{n-1} \geq 0 \;\; \forall\, t \in [0,1]\}$
"Proof": cone: let $c \in K$. then $\theta c \in K$ obviously $(\theta > 0)$.
convex: let c & $d \in K$. To show $\theta c + (1-\theta) d$ is in K, we must have
$\theta c_1 + (1-\theta) d_1 + (\theta c_2 + (1-\theta) d_2) t + \ldots (\theta c_n + (1-\theta) d_n) t^{n-1} \geq 0$
since $= \theta(c_1 + c_2 t \ldots) + (1-\theta)(d_1 + d_2 t \ldots) \geq 0$.
solid: (not sure) claim: interior is where $c_1 + c_2 t + \ldots > 0$

DOUBT — Doesn't the notion of "interior" already require some sort of ordering?

pointed: trivial (show no line)

Now for $c, d \in \mathbb{R}^n$, $c \preceq_K d$ iff $c_1 - d_1 + (c_2 - d_2) t + \ldots > 0 \Leftrightarrow c_1 + c_2 t + \ldots > d_1 + d_2 t + \ldots$

Properties:
(1) $\preceq_K$ is preserved under addition: if $x \preceq_K y$ & $u \preceq_K v \Rightarrow x + u \preceq_K y + v$ (8 #5)
(2) $\preceq_K$ is transitive: if $x \preceq_K y$ & $y \preceq_K z \Rightarrow x \preceq_K z$
(3) $\preceq_K$ is preserved under non-negative scaling: if $x \preceq_K y$ & $\alpha \geq 0 \Rightarrow \alpha x \preceq_K \alpha y$
(4) $\preceq_K$ is reflexive: $x \preceq_K x$
(5) $\preceq_K$ is anti-symmetric: if $x \preceq_K y$ & $y \preceq_K x$ then $x = y$ (8 #5)
(6) $\preceq_K$ is preserved under limits: if $x_i \preceq_K y_i$ for $i = 1, 2, \ldots$ $x_i \to x$ & $y_i \to y$ as $i \to \infty$, then $x \preceq_K y$

Proofs: most are simple, others in rough.

(7) if $x \prec_K y \Rightarrow x \preceq_K y$ (8) if $x \prec_K y$ & $u \preceq_K v \Rightarrow x + u \prec_K v + y$
(9) if $x \prec_K y$ & $\alpha > 0 \Rightarrow \alpha x \prec_K \alpha y$ (10) $x \nprec_K x$ (11) if $x \prec_K y \Rightarrow$ for u & v small enough, $x + u \prec_K y + v$

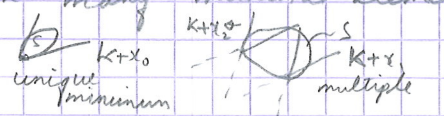
DOUBT $\preceq_K$ instead of $\prec_K$ (for 9)

not sure what "small" would mean (for 11)

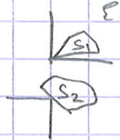
§ 2.4.2 Minimum & minimal elements

Def^n: Linear ordering := any two elements of a set are comparable.
Remark: In reals, either $x \le y$ or $y \le x$. This doesn't hold for generalised inequalities.
Implication: Notions of min & max get complicated.
Def^n: Minimum Element := $x \in S$ is a minimum element if $\forall\, y \in S$, $x \preceq_K y$.
Imp. Def^n: Maximum Element := (implicit)
Def^n: Minimal Element := $x \in S$ is a minimal element if $y \in S$, $y \preceq_K x$ only if $y = x$ ⇐

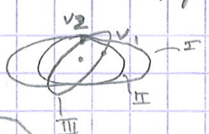
Claim: A set can have many minimal elements but only one minimum element.
Proof: ⌂. e.g.



Alt^n Def^n: Minimum Element := $x \in S$ is minimum iff $S \subseteq x + K$
Alt^n Def^n: Minimal Element := $x \in S$ is minimal iff $(x - K) \cap S = \{x\}$
Remark: For reals, these coincide.



E.g. (1): Consider: $\mathbb{R}^2_+$ (component wise inequality inducing cone). $x \preceq y$ would mean y is on top & right of x. If x is a minimum element then all points of the set lie on top & right of x. If x is a minimal element, then no point of the set lies below & on the left of the set.
(2): S of Symmetric Matrices: $\forall\ A \in S^n_{++}$, consider an ellipsoid (centred at origin)
$$\mathcal{E}_A = \{x \mid x^T A^{-1} x \le 1\}$$
Define: $A \preceq B$ iff $\mathcal{E}_A \subseteq \mathcal{E}_B$ (assumed for now; not worrying about cones)
: $S = \{P \in S^n_{++} \mid v_i^T P^{-1} v_i \le 1, i = 1 \dots k\}$ for $v_1 \dots v_k \in \mathbb{R}^n$ fixed.
Remark: This corresponds to the set of ellipsoids that contain the points $v_1 \dots v_k$
Claim: The set S has no minimum element
Justy^n: ∵ ∃ intersecting ellipsoids ⇒ ∃ ellipsoids that are not comparable. A minimum element must be comparable (smaller) than all elements of a set.
Illustra^n: Minimal element.



I is not minimal ∵ ∃ a "smaller" ellipsoid containing v_1, v_2, II for eg. Similarly for II. III is a minimal element ∵ ∄ (visually) an ellipsoid contained in III (∴ less than III) that encloses v_1 & v_2.

§ 2.5 Separating and Supporting Hyperplanes

§ 2.5.1 Separating hyperplane theorem

Thm: For convex sets C & D, s.t. $C \cap D = \phi$, $\exists\ a \ne 0$ & b s.t. $a^T x \le b\ \forall\, x \in C$ & $a^T x \ge b\ \forall\, x \in D$.
: For (same) C & D, (the affine f^n) $a^T x - b$ is nonpositive on C & nonnegative on D.
Def^n (Impl): Separating Hyperplane := The hyperplane $\{x \mid a^T x = b\}$

Proof: (special case)
Recall: Euclidean Distance := $\mathrm{dist}(C, D) := \inf\{\|u - v\|_2 \mid u \in C, v \in D\}$
Assume: (a) Euclidean distance is positive
(b) ∃ points $c \in C$ $d \in D$ s.t. the minimum is achieved, $\|c - d\|_2 = \mathrm{dist}(C, D)$.
1. Def^n: $a := d - c$, $b := \dfrac{\|d\|_2^2 - \|c\|_2^2}{2}$
objective: Show that the affine f^n $f(x) = a^T x - b = (d - c)^T \left[x - \left(\tfrac{1}{2}\right)(d + c)\right]$ ⌂#6
is non-positive on C & non-negative on D.
remark: This is a hyperplane perpendicular to the line segment between c & d, passing through its midpoint.
"proof start": $f(u) = (d - c)^T \left(u - d + \left(\tfrac{1}{2}\right)(d - c)\right) = (d - c)^T (u - d) + \tfrac{1}{2}\|d - c\|_2^2$
D is non-neg
Suppose $u \in D$ was s.t. $f(u) < 0$ (on the contrary).
⇒ $(d - c)^T (u - d) < 0$.
NB: $\left.\dfrac{d}{dt}\|d + t(u - d) - c\|_2^2\right|_{t=0} = 2\,(d - c)^T (u - d) < 0$ ⌂#6

$\Rightarrow$ for small $t>0$ we must have $\|d + t(u-d) - c\|_2 < \|d-c\|_2$
viz. $d+t(u-d)$ is closer to c than d is.
$\quad = tu + (1-t)d$
Now this point is a convex combination of points u & d so must be in D. But that's a contradiction since d was assumed closest to c in set D. $\square$

E.g. Seperation of an affine & a convex set: assume: (1) C is convex, say, & D is affine, $=\{Fu+g \mid u \in \mathbb{R}^m\}$ where $F \in \mathbb{R}^{n\times m}$. (2) C & D are disjoint
concl^n: By the hyperplane thm $\exists$ an $a \neq 0$ & b s.t. $a^Tx \leq b \ \forall \ x \in C$ & $a^Tx \geq b \ \forall \ x \in D$
explicit plane (constr^n): For $x \in D$, $x = Fu + g \Rightarrow a^TFu \geq b - a^Tg \ \forall \ u \in \mathbb{R}^m$.
A linear f^n is bounded below only when it's zero (in $\mathbb{R}^m$). #7
$\Rightarrow a^TF = 0 \Rightarrow a^Tg \geq b. \Rightarrow b \leq a^Tg$
We obtain a from $a^TF=0$ & use $b = a^Tg$. This specifies one possible plane, $a^Tx = b$.

Def^n: Strict Seperation of C & D

:= (For the aforesaid hyperplane) $a^Tx < b \ \forall \ x \in C$ & $a^Tx > b \ \forall x \in D$

E.g. Strict seperation of a point & a closed convex set.
consider: (1) C is a closed convex set
(2) a point $x_0 \notin C$.
claim: $\exists$ a hyperplane, strictly separating x_0 from C.
proof: consider a set $B := \{x_0 + u \mid \|u\|_2 \leq \epsilon\}$ for some $\epsilon > 0$.
by thm: $\exists \ a \neq 0$ & b s.t. $a^Tx \leq b \quad x \in C$
$a^Tx \geq b \quad x \in B$
This means $a^T(x_0+u) \geq b \quad \forall \ \|u\|_2 \leq \epsilon$. #8
This is minimized by $u = \dfrac{-\epsilon a}{\|a\|_2}$ (see)
$\Rightarrow a^Tx_0 \geq b + \epsilon\|a\|_2 \Rightarrow a^Tx_0 > b'$
where $b' = \dfrac{\epsilon\|a\|_2}{2}$ for instance (anything $< \epsilon\|a\|_2$ really)
Now $a^Tx \leq b$ for $x \in C \Rightarrow a^Tx < b + \dfrac{\epsilon\|a\|_2}{2} = b'$
So we have strict seperation with $a^Tx = b'$ $\left(= b + \dfrac{\epsilon\|a\|_2}{2}\right)$

Quick implic^n: "Proof" that a closed set is the intersection of all halfspaces that contain it.
Assume: C is convex closed
S is the set intersect^n of halfspaces.
NB: If $x \in C$ obviously $x \in S$.
We show: $x \in S \Rightarrow x \in C$.
assume: $x \in S$, but $x \notin C$. By strict seperation, $\exists$ a hyperplane b/w x & C. $\Rightarrow \exists$ a halfspace containing C but not x. $\Rightarrow x \notin S$, contradiction!

Converse separating hyperplane thms

false statement: If $\exists$ a hyperplane (separating) $\Rightarrow C \cap D = \phi$.
counter e.g.: $C = D = \{0\} \subseteq \mathbb{R}^3$. The $x=0$ seperates C & D but $C \cap D = \{0\} \neq \phi$.
(Note, seperates $\neq$ strictly separates)
converse thm: If (1) C & D are convex with C open and (2) $\exists$ a separating hyperplane b/w C & D $\Rightarrow C \cap D = \phi$.
$\quad \Updownarrow$
$\exists$ an affine f^n $f \leq 0$ in C & ≥ 0 in D

"Proof":
combined thm: consider convex sets C & D, with atleast one open, are disjoint $\Leftrightarrow \exists$ a separating hyperplane.

e.g. Theorem of alternatives for strict linear inequalities

target: Find neccessary & sufficient conditions for solvability of $Ax < b$.

NB: f must be <0 in C. $\because$ If $f=0$ at some point, then f would become positive somewhere near that point (along some dir, since it's like a linear f^n, the sign must change if $f=0$ somewhere. no direction is ruled out really because of openness of C (roughly speaking)), which is a contradiction.
This means C & D must be disjoint $\because$ f is negative on C & non-positive on D.

NB: The inequality is unfeasable $\Leftrightarrow$ the (convex) sets $C = \{b - Ax \mid x \in \mathbb{R}^n\}$ & $D = \mathbb{R}^m_{++} = \{y \in \mathbb{R}^m \mid y \succ 0\}$ do not intersect.

NB2: D is open, C is affine

Concl^n: From the thm, C & D are disjoint $\Leftrightarrow$ $\exists$ a seperating hyperplane, i.e. $\exists$ a $\lambda \in \mathbb{R}^m$ & $\mu \in \mathbb{R}$ s.t. $\lambda^T y \le \mu$ on C & $\lambda^T y \geq \mu$ on D.

Simplif^n: The first yields $\lambda^T(b - Ax) \le \mu \;\; \forall x. \Rightarrow A^T\lambda = 0$ (else you can't bound a linear f^n from below) & $\lambda^T b \le \mu$ (#9)

The second yields $\lambda^T y \geq \mu \;\; \forall\; y \succ 0 \Rightarrow$ (see rough) $\mu \le 0$ & $\lambda \succeq 0$, $\lambda \neq 0$.

Putting them together we have

Final Result: The inequality is unfeasable $\Leftrightarrow$ $\exists$ a $\lambda \in \mathbb{R}^m$ s.t. $\boxed{\lambda \neq 0,\ \lambda \succeq 0,\ A^T\lambda = 0,\ \lambda^T b \le 0}$

Remark: This is also a system of linear eq^ns in λ.

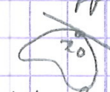
$\therefore$ For a given A & b, exactly one of these sets of equations is solvable.

Nomencl: These form a pair of alternative eq^ns.

§2.5.2 Supporting hyperplanes

Def^n: Boundary := for $C \subseteq \mathbb{R}^n$, $\mathrm{bd}\, C := \mathrm{cl}\, C \setminus \mathrm{int}\, C$.

(Expl) Def^n: Supporting Hyperplane := for $x_0 \in \mathrm{bd}\, C$, for $a \neq 0$, if $a^T x \le a^T x_0 \;\; \forall\, x \in C$ then the hyperplane $\{x \mid a^T x = a^T x_0\}$ is a supporting hyperplane.



Remark: Geometrically, its a tangent to the set.

Thm: Supporting Hyperplane Theorem: For a non-empty convex set C & any point $x_0 \in \mathrm{bd}\, C$, $\exists$ a supporting plane at x_0

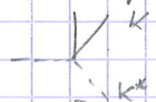
Proof: (1) If int(C) is non empty, then result follows from the hyperplane thm applied to int(C) & $\{x_0\}$.

DOUBT (2) If int(C) is empty $\overset{?}{\Rightarrow}$ C has an affine dimension less than n. Any hyperplane containing this affine set, contains C & x_0. This then is a (trivial) supporting hyperplane.

Inverse Thm: If a set, C, is closed, int(C) $\neq 0$ & it has a supporting hyperplane $\forall$ bd(C) then its convex.

§2.6 Dual cones and generalized inequalities. 19 Oct 2016

§2.6.1 Dual cones



Def^n: Dual Cone := $K^* = \{y \mid x^T y \geq 0 \;\; \forall\; x \in K\}$ where K is a cone.

Claim: K^* is a cone & is always convex (even when K is not)

NB: Geometrically $K^* = \{y \mid -y$ is the normal of a hyperplane that supports K at the origin$\}$

e.g. (1) Subspace: dual cone of a subspace $V \subseteq \mathbb{R}^n$ is its orthonormal complement $V^\perp = \{y \mid v^T y = 0 \;\forall v \in V\}$

(2) Non-negative orthant: (should help both geometrically & algebraically) the cone $\mathbb{R}^n_+$ is its own dual: $x^T y \geq 0 \;\; \forall\; x \succeq 0 \Leftrightarrow y \succeq 0$.

(3) Positive semidefinite cone: Consider S^n & use $\mathrm{tr}(XY)$ inner product.

Claim: S^n_+ is self-dual, i.e. for $X, Y \in S^n$, $\mathrm{tr}(XY) \geq 0 \;\forall X \succeq 0 \Leftrightarrow Y \succeq 0$

Proof: Suppose $Y \notin S^n_+$. $\Rightarrow \exists\; q \in \mathbb{R}^n$ s.t. $q^T Y q = \mathrm{tr}(q q^T Y) < 0$. $Y \notin S^{n*}_+ \Leftarrow$ (#10)

Suppose $X, Y \in S^n_+$. We can write $X = \sum \lambda_i q_i q_i^T$ with $\lambda_i \geq 0$

$\Rightarrow \mathrm{tr}(YX) = \sum \mathrm{tr}(Y \lambda_i q_i q_i^T) = \sum \lambda_i\, q_i^T Y q_i \geq 0 \;\; \because Y \succeq 0$ & $\lambda_i \geq 0$

$\Rightarrow Y \in (S^n_+)^*$

(4) Dual of a norm cone:

Assume: $\|\cdot\|$ is a norm on $\mathbb{R}^n$

Recall: A cone $K := \{(x, t) \in \mathbb{R}^{n+1} \mid \|x\| \le t\}$

Claim: Dual cone $K^* = \{(u, v) \in \mathbb{R}^{n+1} \mid \|u\|_* \le v\}$ where $\|u\|_* = \sup\{u^T x \mid \|x\| \le 1\}$ (the dual norm, see A.1.6 & rough page 3-4)

Proof Idea: Show that

$x^T u + t v \geq 0$ whenever $\|x\| \le t \Leftrightarrow \|u\|_* \le v$

(From the Def^n of a cone, we have $\binom{x}{t}^T (u\ v) \geq 0$ for all points in the domain)

Proof: Start with RHS $\Rightarrow$ LHS; Let $\|u\|_* \le v$, & $\|x\| \le t$ for a $t > 0$ (for $t = 0$ $x = 0$ $\Rightarrow u^T x + v t \geq 0$)

Recall: Defⁿ of dual norm := $\sup\{u^T x \mid \|x\| \leq 1\} =: \|u\|_*$

Using the defⁿ of dual norm & that $\|\frac{-x}{t}\| \leq 1$, we have

$$u^T\left(\frac{-x}{t}\right) \leq \|u\|_* \leq v \quad \Rightarrow \quad u^T x + vt \geq 0$$

Next LHS ⇒ RHS; Let $\|u\|_* > v$ (i.e. RHS doesn't hold).
Now from defⁿ of the dual norm ∃ an x with $\|x\| \leq 1$ s.t. $x^T u > v$. For $t = 1$, we have $u^T(-x) + v < 0$ which violates the LHS (for $t=1$). !RHS ⇒ !LHS
LHS ⇒ RHS.

Claim: Properties of the dual cone (see rough page 3)
(a) K^* is closed & convex (b) $K_1 \subseteq K_2 \Rightarrow K_2^* \subseteq K_1^*$
(c) If K has a non-empty interior, then K^* is pointed
(d) If closure of K is pointed then K^* has non-empty interior
(e) K^{**} is the closure of the convex hull of K.
(Hence if K is convex & closed, $K^{**} = K$)

~ TODO: EX 2.31

§ 2.6.2 Dual Generalised Inequalities

Assume: K is a proper convex cone.
NB: It induces a generalized inequality $\preceq_K$
NB: The dual cone K^* is also proper & convex.
∴ It induces a generalized inequality $\preceq_{K^*}$
Defⁿ: Generalized inequality $\preceq_{K^*}$ is the "dual of" the generalized inequality $\preceq_K$

Properties: (a) $x \preceq_K y \Leftrightarrow \lambda^T x \leq \lambda^T y \ \forall \lambda \succeq_{K^*} 0$
(b) $x \prec_K y \Leftrightarrow \lambda^T x < \lambda^T y \ \forall \lambda \succeq_{K^*} 0, \lambda \neq 0$

NB: Since $K^{**} = K$, dual associated with $\preceq_{K^*}$ is $\preceq_K$; ⇒ properties hold with K^* & K swapped.

E.g. 2.26 Theorem of alternatives for linear strict generalized inequalities
Assume: $K \subseteq \mathbb{R}^m$ is a proper cone.