Convex Optimization:
CVX 101

Ian Quah (itq)

October 30, 2017

Week 2: Convex Sets

Affine sets

- Defn: Contains the line through any two distinct points in a set
- Proof: Affine set contains line through any two points

  Given:

  1) $\{x|Ax = b\}$ (Set constructor)

  2) x = θx_1 + $(1-\theta)x_2$ with $\theta \in \boldsymbol{R}$ (Definition of affine)

  3) Ax_1 = b and Ax_2 = b (Two points chosen arbitrarily)

$$\begin{aligned} Ax &= A(\theta x_1 + (1-\theta)x_2) && \text{(Fact 2)} \\ Ax &= \theta A x_1 + A(1-\theta)x_2 \\ Ax &= \theta b + (1-\theta)b && \text{(Fact 3)} \\ &= b \end{aligned}$$

- Conversely, any affine set can be expressed as solution set of system of linear equations

Convex sets

- Line segment: between x_1 and x_2: all points

$$x = \theta x_1 + (1-\theta)x_2, \text{ with } 0 \leq \theta \leq 1$$

- Convex sets: contains line segment between any two points in the set

$$x_1, x_2 \in \mathcal{C}, 0 \leq \theta \leq 1 \implies \theta x_1 + (1-\theta)x_2 \in \mathcal{C}$$

Convex combination and convex hulls

- Convex combination of $x_1, ..., x_k$: any point x of the form

$$x = \theta_1 x_1 + ... + \theta_k x_k \text{ with}$$

$$1)\theta_1 + ... + \theta_k = 1$$

$$2)\theta_i \geq 0$$

- Convex hull conv S: set of all convex combinations of points in S

  E.g: Consider a square with dots on all 4 sides. The area within the set is the convex hull of the set S (which contains all 4 dots), as we can represent any point *WITHIN* the area as some convex combination of the 4 points (or less, if a θ is 0)

Convex cone

- Conic (nonnegative) combination of x_1, x_2: any point x of the form

$$x = \theta_1 x_1 + \theta_2 x_2 \text{ with}$$

$$\theta_1, \theta_2 \geq 0$$

- Convex cone: set that contains all conic combinations of points in the set
- visually: a pie slice. Lies in a plane even in high dimensions, since we have 3 points

Hyperplanes and halfspaces

- hyperplane: set of the form $\{x|a^T x = b\}$ with a $\neq 0$
- halfspace: set of the form $\{x|a^T x \leq b\}$ with a $\neq 0$ such that a is the normal vector
- hyperplanes are affine and convex, whereas halfspaces are convex
- **Proof: Verifying why the halfspace is convex and NOT affine**

  1) $\{x|a^T x \leq b\}$ (Set constructor)

  2) x = θx_1 + $(1-\theta)x_2$ with $\theta \in \boldsymbol{R}$(Definition of affine)

  3) x = θx_1 + $(1-\theta)x_2$ with $0 \leq \theta \leq 1$ (Definition of convex)

  4) $a^T x_1 \leq$ b and $a^T x_2 \leq$ b (Two points chosen arbitrarily)

$$
\begin{aligned}
\text{WTS: } a^T x &\leq b \\
a^T &= a^T(\theta x_1 + (1-\theta)x_2) && \text{(By definition in 2/3)} \\
&= \theta a^T x_1 + (1-\theta)a^T x_2 \\
&\leq \theta b + (1-\theta)b && \text{(Explanation 1)} \\
&\leq b
\end{aligned}
$$

1) We can put in the inequality, because both are non-negative. At this point, we cannot use 2 anymore, because if we chose any number larger than 1, we would end up with a negative now. Thus, it cannot be affine but is instead convex

Euclidean Ball and Ellipsoids

- Euclidean ball: center x_c and radius r:

$$B(x_c, r) = \{x| \; ||x - x_c||_2 \leq r \tag{1}$$
$$= \{x_c + ru| \; ||u||_2 \leq 1\} \tag{2}$$

1) any point that is within r from the center is contained

2) any point that is some multiple of the radius, where that multiple is between 0 and 1 is contained

- Ellipsoid: set of the form

$$E(x_c, r) = \{x|(x - x_c)^T P^{-1}(x - x_c) \leq 1\} \tag{1}$$
$$= \{x_c + Au| \; ||u||_2 \leq 1\} \tag{2}$$

1) P $\in \boldsymbol{S}^n_{++}$ i.e P symmetric positive definite

2) A is square and non-singular, and can be thought of as a transformation on the unit ball, u

The ellipsoid is a generalization of the Euclidean ball. If P = r^2 $\boldsymbol{I}$, we have that they are equal

Polyhedra or Polytope

- Solution set of finitely many linear inequalities and equalities

  Ax $\prec$ b, Cx = d

$$\text{Where } A \in \boldsymbol{R}^{mxn}, C \in \boldsymbol{R}^{pxn}, \text{ and } \prec \text{ is componentwise inequality}$$

- Is the intersection of finite number of halfspaces and hyperplanes

Positive Semidefinite Cone

- $\boldsymbol{S}^n$ is set of symmetric n x n matrices
- $\boldsymbol{S}^n_+$ = {X $\in \boldsymbol{S}^n | X \succeq 0\}$: positive semidefinite n x n matrices

$$X \in \boldsymbol{S}^n_+ \iff z^T X z \geq 0 \forall z$$

  Note, that a linear combination of

  $\boldsymbol{S}^n_+$ is a convex cone

- $\boldsymbol{S}^n_{++}$ = {X $\in \boldsymbol{S}^n | X \succ 0\}$: positive definite n x n matrices

Operations that preserve Convexity

- Can always work:

$$\text{arbitrary } x_1, x_2 \in \mathcal{C}, 0 \le \theta \le 1 \implies \theta x_1 \text{ and } (1-\theta)x_2 \text{ is in the set}$$

- Show that C is obtained from simple convex sets (hyperplanes, halfspaces, norm balls, ..) by operations that preserve convexity
  - intersection
  - affine functions
  - perspective function
  - linear-fractional functions

Preserving Convexity: Intersection

- Intersection of any number of convex functions is convex

$$\text{eg: } S = \{x \in \boldsymbol{R}^m | abs(p(t)) \le 1 \text{ for } abs(t) \le \frac{\pi}{3}\}$$

  where $p(t) = x_1 \cos t + x_2 \cos 2t + ... + x_m \cos mt$ for m = 2

- Intuition behind example

  Repheasing S:

$$S_t = \{x \in \boldsymbol{R}^m || p(t)| \le 1\}$$

  - S_t is a single point. It is the intersection of two parallel halfspaces (which are linear, as they are a linear combination of cos bla) with opposite orientations (hyperslab)
  - As we modify t, we see that the slab produced by the two parallel lines rotates, and changes width.
  - Thus, S is the intersection of those slabs.

Preserving Convexity: Affine Function

- Suppose f: $\boldsymbol{R}^n \rightarrow \boldsymbol{R}^m$ is affine (f(x) = Ax + b with A $\in \boldsymbol{R}^{mxn}, b \in \boldsymbol{R}^m$)
  - The image of a convex set under f is convex

$$S \subseteq \boldsymbol{R^n} \text{ convex } \implies f(S) = \{f(x)|x \in S\} \text{ convex}$$

  - The inverse image f^{-1}(C) of a convex set under f is convex

$$C \subseteq \boldsymbol{R^m} \text{ convex } \implies f^{-1}(C) = \{x \in \boldsymbol{R^n}|f(x) \in C\} \text{ convex}$$

    This is the case even when the forward function isn't a strict function but more of a relation (many - one). We can still call an inverse on the set

- Examples
  - scaling, translation, projection
  - solution set of linear matrix inequality $\{x|x_1A_1 + ... + x_mA_m \succeq B\}$ with A_i, B $\in \boldsymbol{S}^P$
    This is basically a linear matrix inequality
  - hyperbolic cone $\{x|x^TPx \leq (c^Tx)^2, c^Tx \geq 0\}$ (with P $\in \boldsymbol{S}_+^n$)

Preserving Convexity: Perspective and linear-fractional function

- perspective function P: $\boldsymbol{R}^{n+1} \rightarrow R^n$:

$$P(x,t) = x/t, \mathbf{dom}P = \{(x,t)|t > 0\}$$

  images and inverse images of convex sets under perspective are convex
  - Basically, divide the first n elements of the set by the last element
  - perspective function: has the effect of normalizing by the last component

- linear-fractional function: f:$R^n \rightarrow R^m$:

$$f(x) = \frac{Ax+b}{c^Tx+d}, \mathbf{dom}f = \{x|c^Tx + d > 0\}$$

  images and inverse images of convex sets under linear-fractional functions are convex
  Note the following:
  - If you compose the perspective function with an affine function you get the linear-fractional
  - denominator is a scalar
  - both numerator and denominator are both affine
  - Can be thought of as c̈onditioningïn probability. We're normalizing the computed numerator

The perspective function says that if the object we see in 3d is convex, then the image in 2d is also convex

Generalized Inequalities

- A Convex cone K $\subseteq \boldsymbol{R}^n$ is a **proper cone** if
  - K is closed (contains its boundary)
  - K is solid (nonempty interior)
  - K is pointed (contains no line)
- Examples:
  - nonnegative orthant K = $\boldsymbol{R}^n_+ = \{x \in \boldsymbol{R}^n | x_i \geq 0, i = 1, ..., n\}$
  - positive semidefinite cone K = $\boldsymbol{S}^n_+$
  - nonnegative polynonmials on [0, 1]:

$$K = \{x \in \boldsymbol{R}^n \ | x_1 + x_2 t + x_3 t^2 + ... + x_n t^{n-1} \geq 0 \text{ for } t \in [0, 1]$$

- Generalized Inequality defined by a proper cone K:

$$x \preceq_K y \iff y - x \in K, x \prec_K y \iff y - x \in \mathbf{int} K$$

  , e.g x $\preceq$ y and X $\prec$ Y (for vectors and matrices)

  Note the above describes inequality and strict inequality

  Examples
  - componentwise inequality (K = $\boldsymbol{R}^n_+$)

$$x \preceq_{\boldsymbol{R}^n_+} y \iff x_i \leq y_i, i = 1, ..., n$$

  - matrix inequality (K = $\boldsymbol{S}^n_+$)

$$X \preceq_{\boldsymbol{S}^n_+} Y \iff Y - X positive semidefinite$$

  - The above examples are so common that we drop the subscript in $\preceq_K$
- **Properties**: many properties of $\preceq_K$ are similar to $\leq$ on $\boldsymbol{R}$, e.g

$$x \preceq_K y, u \preceq_K v \implies x + u \preceq_K y + v$$

Minimum and Minimal Elements

- $\preceq_K$ is not in general a linear ordering: we can have x $\not\preceq_K$ y and y $\not\preceq_K$ x (can say that they are incomparable: e.g two vectors in $\boldsymbol{R}^2$: [2,1] and [1, 2])
- When moving from a total ordering into a partial ordering we have interesting things that happen:
- x $\in S$ is the **minimum element** of S wrt $\preceq_K$ IF

$$y \in S \implies x \preceq_K y$$

  Means that x is the minimum if 1) comparable to every other element in the set, AND it is smaller than or equal to all of them (and is unique)
- x $\in S$ is a **minimal element** of S wrt $\preceq_K$ IF

$$y \in S, y \preceq_K x \implies y = x$$

  If you're in the set and you're comparable to it, then you cannot be less than x without being less than it

Separating Hyperplane Theorem

- **Separating hyperplane theorem**

  If C and D are disjoint convex sets, then there exists a $\neq$ 0, b such that

$$a^T x \leq b for x \in C, a^T x \geq b for x \in D$$

  , such that the hyperplane $\{x|a^T x = b\}$ separates C and D

  strict separation requires additional assumptions (e.g. C is closed and D is a singleton)
- **Supporting Hyperplane Theorem**

  Defn: If C is convex, then $\exists$ supporting hyperplane at every boundary point of C

  supporting hyperplane to set C at boundary point x_0:

$$\{x|a^T x = a^T x_0\}$$

  where a$\neq$0 and a^Tx $\leq$ $a^T x_0$ for all x $\in$ C

Dual Cones and generalized inequalities

- **Dual cone** of a cone K:

$$K^* = \{y | y^T x \geq 0 \forall x \in K\}$$

Things that make a non-negative inner-product with K. K^{**} == K.

  - K = $\boldsymbol{R}_+^n : K^* = \boldsymbol{R}_+^n$
    If you draw it out, you find that you get something that is 90°
    Vector is nonnegative IFF forms nonnegative inner product with all nonnegative vectors
  - K = $\boldsymbol{S}_+^n : K^* = \boldsymbol{S}_+^n$
    Basically saying that given some symmetric matrices, X, Y,

$$trace(X, Y) = \sum_{i,j} X_{ij} Y_{ij} == vec(X)^T vec(Y)$$

    Matrix is PSD IFF its trace against all PSD matrices is nonnegative
  - K = $\{(x, y) | ||x||_2 \leq t\} : K^* = \{(x, y) | ||x||_2 \leq t\}$
  - K = $\{(x, y) | ||x||_1 \leq t\} : K^* = \{(x, y) | ||x||_\infty \leq t\}$

- first 3 examples are **self-dual** cones

  dual cones of proper cones are proper, hence define generalized inequalities:

$$y \succeq_{K^*} 0 \iff y^T x \geq 0 \forall x \succeq_K 0$$