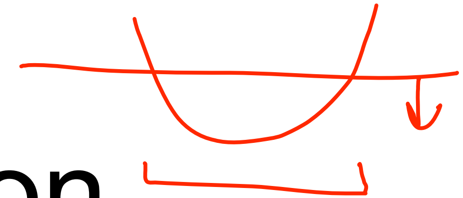
Convex optimization



minimize $f(x)$ subject to $Ax = b$

($x \in \mathbb{R}^n$, f convex, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$)

$g_i(x) \le 0 \quad \forall i \in I$ (g_i convex)

e.g., min $x^2 + y^2$ s.t. $\underbrace{1 - x - y}_{g_i(x,y)} \le 0$

Linear inequalities: OK

Positivity: OK

(just include in g_i's)

Optimality

- Optimal value:

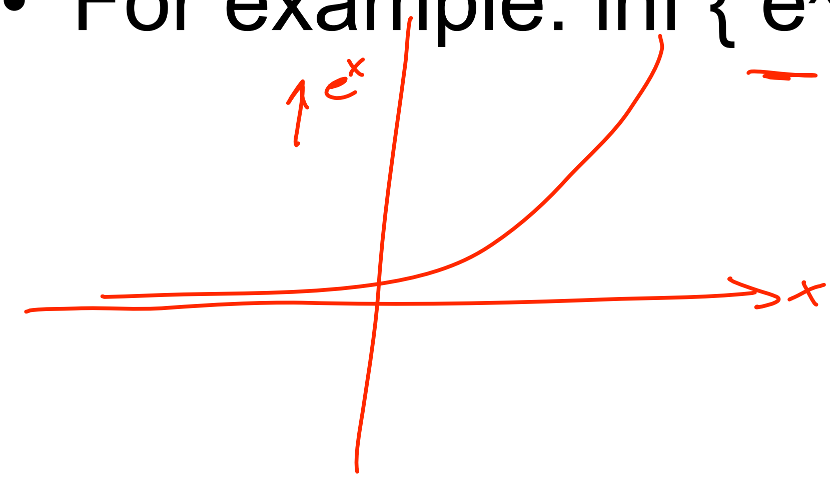
  $v^* = \inf \{ f(x) \mid g_i(x) \leq 0 \ (\forall i), Ax = b \}$

  Also, $v^* \leq x$ $\forall x \in S$

- Definition of $v^* = \inf S$: $\subseteq \mathbb{R}$ → infimum

  $\Longleftrightarrow \forall \epsilon > 0 \ \exists x_\epsilon \in S. \ v^* + \epsilon \geq x_\epsilon$

- For example: $\inf \{ e^x \mid x \text{ real} \} =$



sup (supremum)

tightest upper bound

Optimal point

- x^* is optimal iff:

- Is there always an optimal point?
  - ex. 1:

  - ex. 2:

Local optima

- min f(x) s.t. $g(x) \leq 0$
- x_0 is a local optimum iff

- In a convex program,

An optimality criterion

- Suppose f(x) is differentiable, convex
- Then x* = arg min f(x) iff
- What about x* = arg min f(x) s.t. $x \in C$?

Proof of criterion

Types of convex program

- Linear program, quadratic program
- Second-order cone program
- Semidefinite program

Example: logistic regression

- Data $x_i \in R$, $y_i \in \{0,1\}$
- $P(y_i = 1 \mid x_i, w, b) = s(w^T x_i + b)$

  $s(z) = 1/(1+\exp(-z))$

- $\max_{wb} P(w, b)\, P(y \mid x, w, b) =$

  $\max_{wb} P(w, b) \prod_i P(y_i \mid x_i, w, b) =$

Ex: minimize top eigenvalue

- Suppose $A_i \in S^{n*n}$ for i = {1, 2, …}
- $\min_{\alpha,A} \ \lambda_1(A)$ s.t.

  $A = \alpha_1 A_1 + \alpha_2 A_2 + \ldots$

Ex: minimum volume ellipsoid

- Given points $x_1, x_2, \ldots, x_k$
- $\min_{A, x_C} \ \text{vol}(A)$ s.t.

  $(x_i - x_C)^T A (x_i - x_C) \le 1 \quad i = 1, \ldots, k$

  $A \in S^{n*n}$

  $A \succcurlyeq 0$

Schur complement

- Symmetric block matrix M =
- Schur complement is S =
- $M \succcurlyeq 0$ iff

Back to min-volume ellipsoid

- $\max_{A, x_C} \log |A|$ s.t.
  $(x_i - x_C)^T A (x_i - x_C) \leq 1 \quad i = 1, \ldots, k$
  $A = A^T,\ A \succcurlyeq 0$
- $\max_{A, B, x_C, z} \log |A|$ s.t.

Ex: soap bubbles

- Q: Dip a bent paper clip into soapy water. What shape film will it make?
- A:
- Write h(x,y) for height
- Suppose (x,y) in S
- Area is:

Soap bubbles

Ex: manifold learning

- Given points $x_1, \ldots, x_m$
- Find points $y_1, \ldots, y_m$
- Preserving **local geometry**
  - neighbor edges N
  - distances
- If we preserve distances

  we also preserve angles

Step 1: “embed” R^n into R^n

- While preserving local distances, move points to make manifold as flat as possible
- max

  s.t.

Step 2: reduce to R^d

- Now that manifold is flat, just use PCA:

Maximizing variance

- max s.t.

- max s.t.

Summary

- Solve SDP to “embed” R^n into R^n
- Use PCA to embed R^n into R^d
- Called “semidefinite embedding” or “maximum variance unfolding”

- Problems?

Problem: solving SDP

- Kernel matrix K
- Idea: suppose we know a subspace that preserves geometry

Ex: convex games

- “Robin hood” game:
  - shoot two arrows at target simultaneously
  - Robin Hood’s cost:
  - Sheriff of Nottingham’s reward:

Convex games: definition

- Players p_1, p_2, …, p_k
- p_i chooses x_i from convex set X_i

  all simultaneously
- Cost to p_i is $f_i(x_1, x_2, \ldots, x_k)$
- Cost f_i is convex in x_i for fixed x_{-i}
- Zero-sum:

  2 players, $f_1(x_1, x_2) = -f_2(x_1, x_2)$

Equilibrium

- Minimax equilibrium: find
  - value v
  - distribution P(x)
- such that
  -
  -

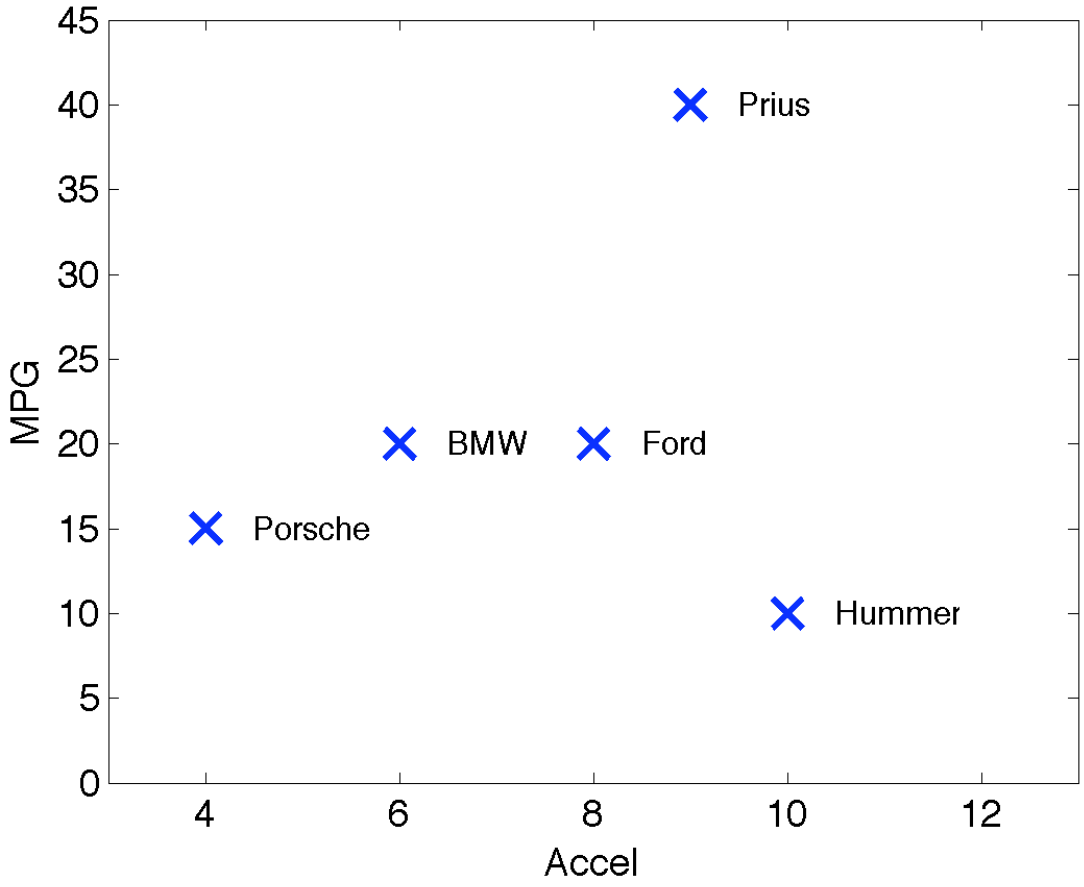
Multi-criterion optimization

- Same feasible region as ordinary CP
- Indecisive optimizer: wants all of

Multi-criterion example

- Buying the perfect car

Pareto optimality



- Pareto optimal =