Linear feasibility problem

min $c^T x$ s.t.

$Ax + b \geq 0$

max $-b^T y$ s.t.

$A^T y = c$, $y \geq 0$

find x s.t.

$Ax + b \geq 0$

Separation oracle

Ellipsoid preview

Difficulties

- How do we get bounding sphere?
- How do we know when to stop?
- Bound region gets complicated–how do we find its center?

Bounding a half-ellipsoid

- General ellipsoid w/ center x_C, shape A:

- Halfspace: $p^T x \le p^T x_c$
- Translate to origin, scale to be spherical

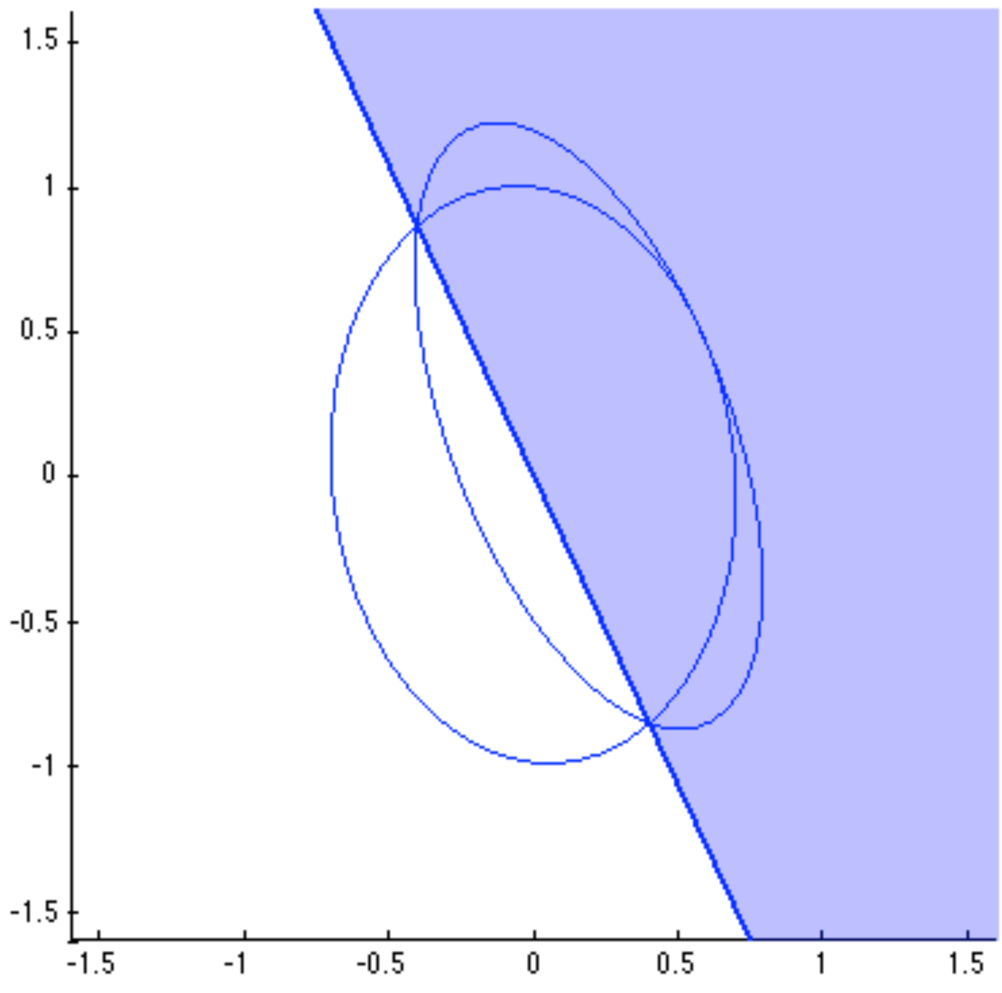
  y = x =

Bounding a half-sphere

- Rotate so hyperplane is axis-normal

- New center z_C:
- New shape B:

For example



Ellipsoid algorithm

- Want to find x s.t. $Ax+b+\eta \geq 0$
- Pick E_0 s.t. $x^* \in E_0$
- for t := 1, 2, …
  - x_t := center of E_t
  - ask whether $Ax_t + b + \eta \geq 0$
    - yes: declare feasible!
    - no: get new constraint w/ normal p_t
  - $E_{t+1} := \text{bound}(E_t \cap \{ x \mid p_t^T x \leq p_t^T x_t \})$
  - if $\text{vol}(E_{t+1}) \leq \varepsilon \text{vol}(E_0)$: declare infeasible!

Getting bounds

- How big does E_0 need to be?
- What should η be?
- How small does ε need to be?

Dotting i's, crossing t's

- What if LF was unbounded?

- What about numerical precision?

Comparison to constraint generation

- Ellipsoid is polynomial, but slow
- Constraint generation has no non-trivial runtime bound, but often much faster

Other algorithms

- Interior point: polynomial, can be very fast
- Simplex: exponential in worst case, but often fast in practice
- Randomized simplex: polynomial [Kelner & Spielman, 2006]
- Subgradient descent: weakly polynomial, but really simple, and fast for some purposes

What’s a subgradient?

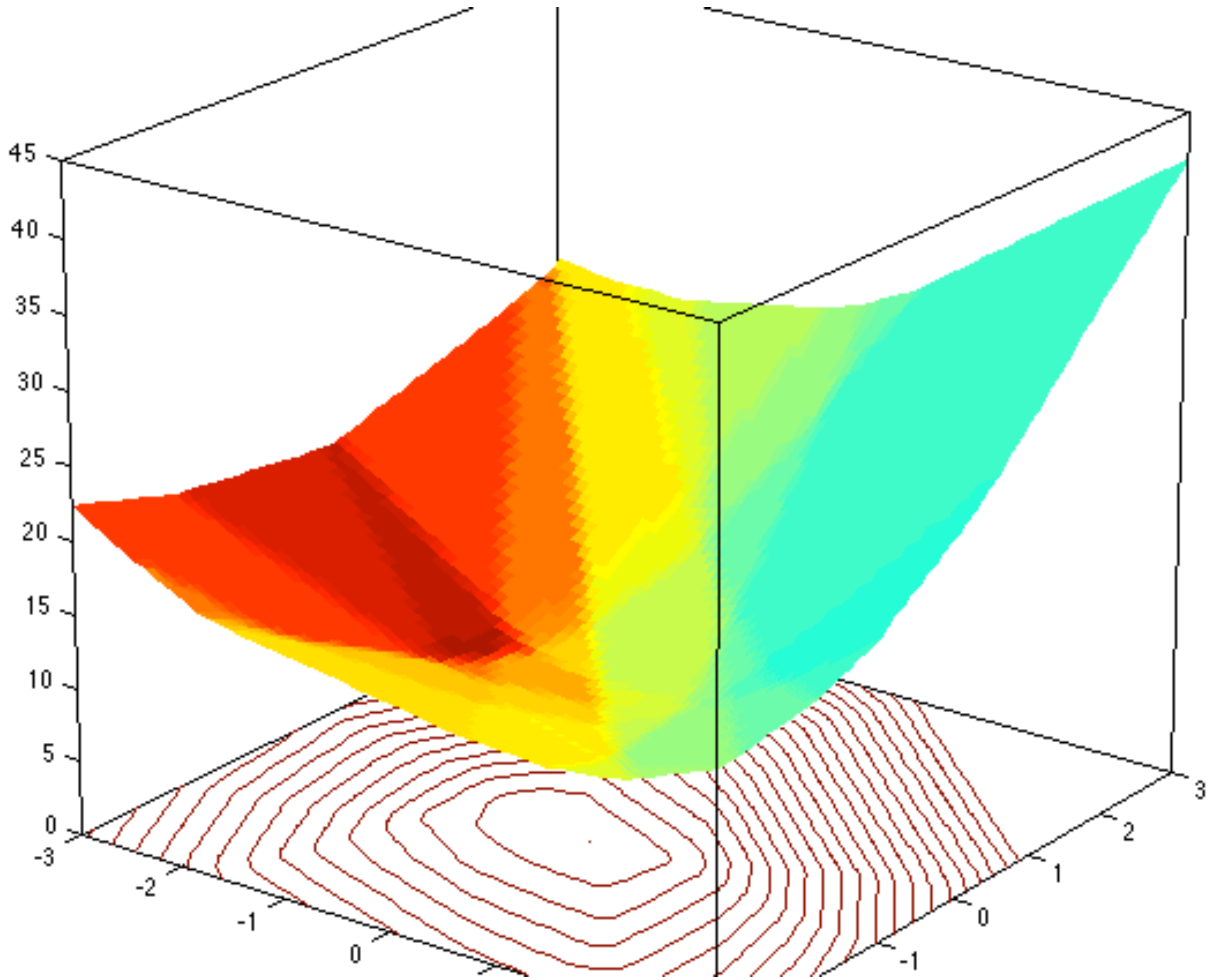
Subgradients for SVMs

- $\min_{s,w,b}$ $||w||^2 + C\sum_i s_i$ s.t.
  $y_i(x_i^T w - b) \geq 1 - s_i$
  $s_i \geq 0$
- Equivalently,

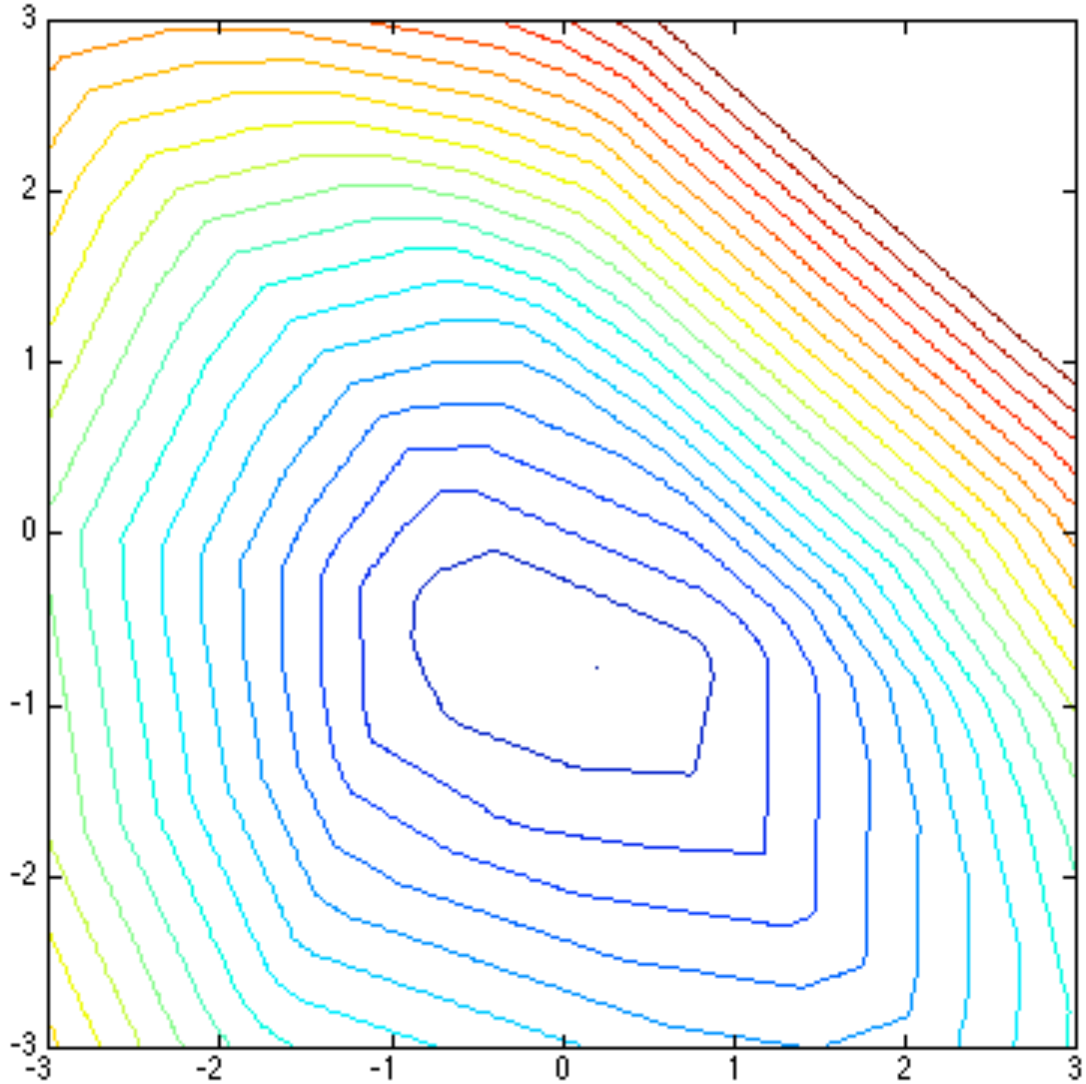
Subgradients for SVMs

- $\min_w L(w) = ||w||^2 + (C/m)\Sigma_i h(y_i x_i^T w)$
- Subgradient of h(z):
- Subgradient of L(w) wrt w:

SVM loss



Subgradient descent



Subgradient descent

- While not tired:

  $g_t = \quad \partial f(x_t)$

  $x_{t+1} = x_t - \eta_t g_t$

Subgradient questions

- How to choose learning rate?
- How to decide when we're tired?
- How to estimate $\partial f(x_t)$?