Abstraction + optimization algorithms

Data / Application

Classify? Regress? Dim. red?

Model

- features
- decision fns
- decision
- over/under-fitting

Optimization problem

- model → something we can train

Optimization algorithms

- Gradient decent
- simplex

Optimization problems

Unconstrained optimization:

- objective function $f(x)$, smooth
- find w^* that minimizes f
- f has global & local minima

- $f(x)$ is convex if $\forall x, y \in \mathbb{R}^d$ a line segment $(x, f(x))$ to $(y, f(y))$ never goes below $f(\cdot)$
- Continuous cnvx. fn. has one of:
  - no minimum
  - one unique global minimum (strong convexity)
  - connected set of local minima w/ equal $f(\cdot)$

Sum of f_i, f_i convex $\forall i$, still convex

Algos for general smooth f:

- Gradient decent
  - ↳ blind (w/ learning rate)
  - ↳ stochastic blind
  - ↳ w/ line search
- Newton's Method (requires Hessian)
- Non linear Conjugate gradient

Algos for non-smooth f:

- gradient decent
- BFGS (Broyden-Fletchel- Goldfarb-Shanno)

line search

- project down to 1D and use that result to inform your step size
  - ↳ in 1D its easy to use secant method / Newton's method
  - ↳ direct search

Constrained optimization

given $f(x)$ and constraints $g(x)=0$

↳ use Lagrangian to make it unconstrained

LP's

$$\min_x \; c^T x \quad : \quad Ax \leq b$$

$x \in \mathbb{R}^d$ | $A \in \mathbb{R}^{n \times d}$ | $b \in \mathbb{R}^n$ | $c \in \mathbb{R}^d$

- optima lie on verticies (when some constraints hold w/ equality)
  are "active"

Hard-margin SVM: support vectors are active constraints

Algos:

- simplex: walk verticies to find optimum
- interior point methods

QPs

$$\min_x \; x^T Q x + c^T x \quad : \quad Ax \leq b$$

$Q \succ 0$ ex. SVM $\min_w \|w\|^2 \leftarrow w^T I w$ $I \succeq 0$

Algos:

- simplex (good in general)
- Sequential minimal optimization (SMO)
- Coordinate descent