Our goal is to solve

$$\min_{x \in K} \widehat{\theta}^T x$$

where K is convex and compact.

1 Interior point methods

Path-following interior point is the following:

- While $t > \varepsilon/m$,
  - "Approximately" solve
    $$\widehat{x}_t \approx \mathrm{argmin}_{x \in K} \underbrace{x^T \widehat{\theta} + t\varphi(x)}_{\Phi_t(x)}$$
    where φ is self-concordant barrier function.
  - $t \hookleftarrow t(1-\gamma)$.

Examples for φ:

- $K = \{x : Ax < b\}$. Then $\varphi_K(x) = \sum_i -\ln(b_i - A_i x)$ works.

Requirements:

1. $\varphi \to \infty$ as $x \to \partial K$.
2. $\nabla_x^3 \varphi[h,h,h] \le 2(\nabla_x^2 \varphi[h,h])^{\frac{3}{2}}$
3. $\nabla_x \varphi[h] \le \sqrt{\nu \nabla_x^2 \varphi[h,h]}$.

Approximate algorithm: Newton's method

$$\widehat{x}^+ \hookleftarrow \widehat{x} - c\nabla^{-2}\varphi(\widehat{x})\nabla\Phi_t(\widehat{x}).$$

Analysis relies on

$$\lambda_t(x) = \sqrt{\nabla\Phi_t(x)\nabla^{-2}\varphi(x)\nabla\Phi_t(x)}$$

being small. I want $\nabla\Phi_t$ to be small measured in norm of Hessian.

$$t^+ := t\left(1 - \frac{d}{\sqrt{\nu}}\right) \tag{1}$$

$$\lambda_{t^+}(\widehat{x}) \le \lambda_t(\widehat{x})(1+c) + c \tag{2}$$

$$\lambda_{t^+}(\widehat{x}^+) \le 2\lambda_{t^+}^2(\widehat{x}). \tag{3}$$

λ doesn't increase too much when you increase the temperature a little; then after doing a Newton step it decreases quickly.

Once $t < \varepsilon$ you have an ε-approximate solution.

We need $(1 - \frac{d}{\sqrt{\nu}})^L < \varepsilon$, so need $K = \Theta(\sqrt{\nu}\ln\left(\frac{1}{\varepsilon}\right))$ epochs.

Does there exist φ for every compact convex K? Nesterov and Nemivroski showed yes, for $\nu = O(n)$. It is constructive but not efficiently computable (volume of polar cones...)

2 Simulated annealing

- While $t > \varepsilon$,
  - Collect N (approximate) samples $X_t^1, \ldots, X_t^N \sim P_{\theta/t}$ where $P_{\theta t(x)} \propto \exp(-\frac{\theta^T x}{t})$ on K.
  - Update $t \hookleftarrow t(1 - \frac{1}{\sqrt{n}})$.
  - Use samples $X_t^1, \ldots, X_t^n$ to "warm start" on new point.

For Hit-and-Run you need warm start and the covariance matrix. Take $N \approx n$ to get good estimate on covariance matrix (isotropically chose).

Hit-and-run:

- Input: Σ, X_0 warm start, P_0 (oracle access)
- For $k = 1, 2, \ldots, K$,
  - Sample $U \sim N(0, \Sigma)$
  - Compute ray $R = \{X_{k-1} + \alpha U : \alpha \in \mathbb{R}\}$.
  - Sample $X_k \sim P_\theta|_R$.
- Return X_K.

Here I only need to be able to check whether points are in the set!

Lovasz, Vempala, etc.: HAR(Σ, X_0, P_0) returns X such that X is approximately distributed according to P_θ, given

1. X_0 is sampled from "nearby" distribution
2. Σ is "almost" Cov(P_θ).
3. $K = cn^3$.

Theorem 2.1. *If $X \sim P_{\theta/t}$, then $\mathbb{E}[\widehat{\theta}^T x] \leq \min_\alpha \widehat{\theta} x + tn$.*

Abernethy-Hazan 2016: There exists a strongly convex barrier function φ such that

$$\text{argmin}_{x\in K} \widehat{\theta}^T \varphi(x) + t\varphi(x) = \mathop{\mathbb{E}}_{x\sim P_{\theta/t}}[x] \tag{4}$$

What is φ? For exponential families

$$A(\theta) = \ln \int_K \exp(-\theta^T x)\, dx \tag{5}$$

$$\varphi(\theta) = A^*(\theta) = \sup_\theta \theta^T x - A(\theta), \tag{6}$$

Fenchel conjugate of A,

$$P_\theta(x) = \exp(-\theta^T x - A(\theta)) \tag{7}$$

$$\nabla A(\theta) = \mathop{\mathbb{E}}_{X\sim P_\theta}[X]. \tag{8}$$

(Check signs.)

Fact:

$$\nabla A(\theta) = \mathrm{argmax}_{x\in K}\, x^T\theta - A^*(x).$$

The last 2 equations explain why this works.

Q: What's the barrier parameter that this gives us? It is n all the time, but this is misleading: you can make parameter better by scaling.

HAR gives the following entropic barrier for PSD cone:

$$\varphi(X) = n\ln(\det X).$$

(We can divide this by n.)

Kalai-Vempala give time $n^{4.5}$, we show time $n^4\sqrt{\nu}$.

Cf. Narayanan—Dikin walk—single step every round, compute covariance at each step.