Convex Optimization

Prof. Nati Srebro

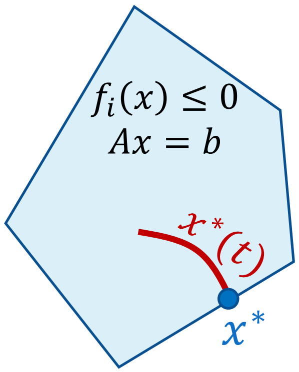
Lecture 12:

Interior Point Methods

Read: Boyd and Vandenberghe Chapter 11

Log Barrier

$$\boxed{\begin{array}{ll} \min_{x \in \mathbb{R}^n} & f_0(x) - \frac{1}{t}\sum_{i=1}^m \log\left(-f_i(x)\right) \\ s.t. & Ax = b \end{array}}$$



- Central Path: $\{x^*(t) | t > 0\}$

- $x^*(t)$ strictly feasible for original (constrained) problem
- We will show: $x^*(t) \rightarrow x^*$

- How suboptimal (for original problem) is $x^*(t)$?
- How do we set t?
- What's the complexity of optimizing with that t?

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f_0(x) - \frac{1}{t}\sum_{i=1}^{m} \log(-f_i(x)) \\ s.t. \quad & Ax = b \end{aligned}$$

Optimum x_t^*, dual opt ν_t^*

$$L_t(x,\nu) = f_0(x) - \frac{1}{t}\sum_i \log(-f_i(x)) + \langle \nu, Ax - b\rangle$$

$$0 = \nabla_x L_t(x_t^*, \nu_t^*) = \nabla f_0(x_t^*) + \sum_i \underbrace{\frac{-1}{t f_i(x_t^*)}} \nabla f_i(x_t^*) + A^\top \nu_t^*$$

Define $\lambda_{t,i}^* = \frac{-1}{t f_i(x_t^*)} > 0$

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f_0(x) \\ s.t. \quad & f_i(x) \le 0,\ Ax = b \end{aligned}$$

x_t^* is strictly feasible

How suboptimal is x_t^* ?

$$L(x,\lambda,\nu) = f_0(x) + \sum_i \lambda_i f_i(x) + \langle \nu, Ax - b\rangle$$

$$\nabla_x L(x_t^*, \lambda\ , \nu_t^*) = \nabla f_0(x_t^*) + \sum_i\ \lambda_i\ \nabla f_i(x_t^*) + A^\top \nu_t^*$$

$$\min_{x \in \mathbb{R}^n} \quad f_0(x) - \frac{1}{t}\sum_{i=1}^{m} \log\big(-f_i(x)\big)$$
$$s.t. \quad Ax = b$$

Optimum x_t^*, dual opt ν_t^*

$$L_t(x, \nu) = f_0(x) - \frac{1}{t}\sum_i \log\big(-f_i(x)\big) + \langle \nu, Ax - b\rangle$$

$$0 = \nabla_x L_t(x_t^*, \nu_t^*) = \nabla f_0(x_t^*) + \sum_i \underbrace{\frac{-1}{t f_i(x_t^*)}} \nabla f_i(x_t^*) + A^\top \nu_t^*$$

Define $\lambda_{t,i}^* = \frac{-1}{t f_i(x_t^*)} > 0$

$$\min_{x \in \mathbb{R}^n} \quad f_0(x)$$
$$s.t. \quad f_i(x) \leq 0,\ Ax = b$$

x_t^* is strictly feasible

How suboptimal is x_t^* ?

$$L(x, \lambda, \nu) = f_0(x) + \sum_i \lambda_i f_i(x) + \langle \nu, Ax - b\rangle$$

Relaxed KKT

$$\nabla_x L(x_t^*, \lambda_t^*, \nu_t^*) = \nabla f_0(x_t^*) + \sum_i \lambda_{t,i}^* \nabla f_i(x_t^*) + A^\top \nu_t^*$$

$$Ax_t^* = b$$
$$f_i(x_t^*) < 0$$
$$\lambda_t^* > 0$$

But:

$$f_i(x_t^*) \cdot \lambda_{t,i}^* = f_i(x_t^*) \cdot \frac{-1}{t f_i(x_t^*)} = -\frac{1}{t}$$

$$\boxed{\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & f_0(x) - \frac{1}{t}\sum_{i=1}^m \log\big(-f_i(x)\big) \\ s.t. & Ax = b \end{array}}$$

Optimum x_t^*, dual opt ν_t^*

$$L_t(x,\nu) = \\ f_0(x) - \frac{1}{t}\sum_i \log\big(-f_i(x)\big) + \langle \nu, Ax - b\rangle$$

$$0 = \nabla_x L_t(x_t^*, \nu_t^*) = \\ \nabla f_0(x_t^*) + \sum_i \underbrace{\frac{-1}{tf_i(x_t^*)}} \nabla f_i(x_t^*) + A^\top \nu_t^*$$

Define $\lambda_{t,i}^* = \frac{-1}{tf_i(x_t^*)} > 0$

$$\boxed{\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & f_0(x) \\ s.t. & f_i(x) \le 0,\ Ax = b \end{array}}$$

x_t^* is strictly feasible

How suboptimal is x_t^* ?

$$L(x,\lambda,\nu) = \\ f_0(x) + \sum_i \lambda_i f_i(x) + \langle \nu, Ax - b\rangle$$

$$\nabla_x L(x_t^*, \lambda_t^*, \nu_t^*) \\ = \nabla f_0(x_t^*) + \sum_i \lambda_{t,i}^* \nabla f_i(x_t^*) + A^\top \nu_t^*$$

$$g(\lambda_t^*, \nu_t^*) = \inf_x L(x, \lambda_t^*, \nu_t^*) = L(x_t^*, \lambda_t^*, \nu_t^*) \\ = f_0(x_t^*) + \sum \frac{-1}{tf_i(x_t^*)} f_i(x_t^*) + \cancel{\langle \nu_t^*, Ax_t^* - b\rangle} \\ = f_0(x_t^*) - \frac{m}{t}$$

λ_t^*, ν_t^* dual (strictly) feasible with

$$f_0(x_t^*) - g(\lambda_t^*, \nu_t^*) = \frac{m}{t}$$

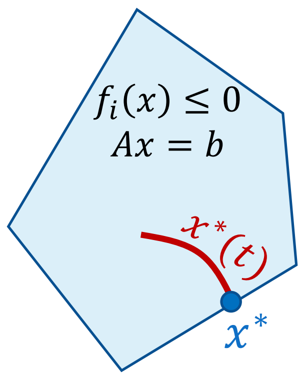
$$\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & f_0(x) - \frac{1}{t}\sum_{i=1}^m \log\big(-f_i(x)\big) \\ s.t. & Ax = b \end{array}$$

Optimum x_t^*, dual opt ν_t^*

$$\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & f_0(x) \\ s.t. & f_i(x) \le 0,\ Ax = b \end{array}$$

x_t^* is strictly feasible

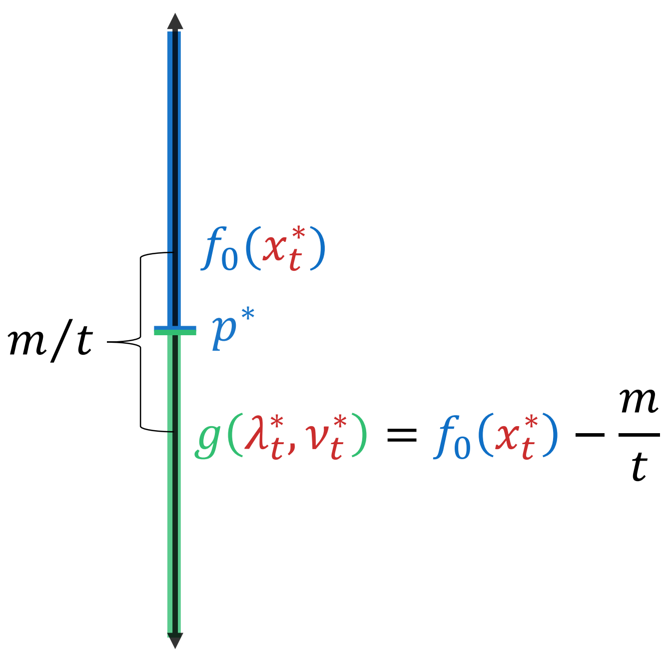
How suboptimal is x_t^* ?



$$L(x,\lambda,\nu) = f_0(x) + \sum_i \lambda_i f_i(x) + \langle \nu, Ax - b\rangle$$

$$\nabla_x L(x_t^*, \lambda_t^*, \nu_t^*) = \nabla f_0(x_t^*) + \sum_i \lambda_{t,i}^* \nabla f_i(x_t^*) + A^\top \nu_t^*$$

$$g(\lambda_t^*, \nu_t^*) = \inf_x L(x, \lambda_t^*, \nu_t^*) = L(x_t^*, \lambda_t^*, \nu_t^*)$$
$$= f_0(x_t^*) + \sum \frac{-1}{t f_i(x_t^*)} f_i(x_t^*) + \cancel{\langle \nu_t^*, Ax_t^* - b\rangle}$$
$$= f_0(x_t^*) - \frac{m}{t}$$



λ_t^*, ν_t^* dual (strictly) feasible with

$$f_0(x_t^*) - g(\lambda_t^*, \nu_t^*) = \frac{m}{t}$$

Log Barrier: Sub-Optimality

$$\boxed{\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & f_0(x) - \frac{1}{t}\sum_{i=1}^m \log\big(-f_i(x)\big) \\ s.t. & Ax = b \end{array}}$$

- $x^*(t)$ is $\frac{m}{t}$-suboptimal for original problem
- To get ϵ-suboptimal, set $t = {}^m/_\epsilon$ and solve using Newton
- Runtime? Number of Newton iterations?
- Assumptions: - $f_0(x)$ is self concordant

  - $f_i(x)$ is quadratic (or linear) ➔ $-\log\big(-f_i(x)\big)$ self conc
- Objective *not* self-concordant, but scaled by t it is:

$$\tilde{f}(x) = tf_0(x) - \sum_{i=1}^m \log\big(-f_i(x)\big)$$

- #Newton iterations:

$$k \propto \left(\tilde{f}\left(x^{(0)}\right) - \tilde{f}\left(x_t^*\right)\right) = t\left(f_0\left(x^{(0)}\right) - f_0\left(x_t^*\right)\right) + \sum_i \log\frac{f_i(x_t^*)}{f_i\left(x^{(0)}\right)} \propto \frac{m}{\epsilon}(\cdots) + (\cdots)$$

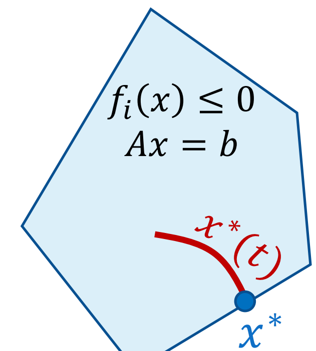
- Overall runtime: $\approx O\left(\frac{m(n+p)^3}{\epsilon}\left(f_0\left(x^{(0)}\right) - f_0\left(x^*\right)\right)\right)$

typically ≤ 0, if $x^{(0)}$ not too close to boundary

Central Path Log Barrier Method

Init:	Feasible $x^{(0)}$ and some $t^{(0)}$
Do:	Solve $t^{(k)}$-barrier problem using Newton starting at $x^{(k)}$
	$x^{(k+1)} \leftarrow x^*(t^{(k)})$
	Stop if $\frac{m}{t} \le \epsilon$
	$t^{(k+1)} \leftarrow \mu \cdot t^{(k)}$ (for some parameter $\mu > 1$)

- When do we stop Newton? Why nested loops OK?
  - Stop at machine precision; OK since Newton gets there quickly
- Sketch of analysis:



$$\overbrace{\frac{\log \frac{m}{\epsilon}}{\log \mu}}^{\#\text{steps to reach } \frac{m}{t} \le \epsilon} \cdot \underbrace{\overbrace{\left(t\left(f_0(x^*_{t/\mu}) - f_0(x^*_t)\right) + \sum_i \log \frac{f_i(x^*_t)}{f_i(x^*_{t/\mu})} \right)}^{\#\text{Newton to optimize barrier problem}}}_{\le\; m(\mu - 1 - \log \mu)} = O\left(\sqrt{m} \log \frac{m}{\epsilon}\right)$$

Using

$\mu = \left(1 + {}^{1}/_{\sqrt{m}}\right)$

Central Path Log Barrier Method

Init: Feasible $x^{(0)}$ and some $t^{(0)}$

Do: Solve $t^{(k)}$-barrier problem using Newton starting at $x^{(k)}$

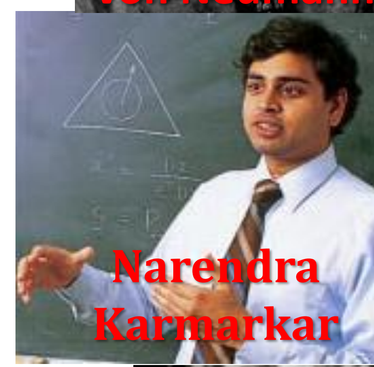
$x^{(k+1)} \leftarrow x^*(t^{(k)})$

Stop if $\frac{m}{t} \leq \epsilon$

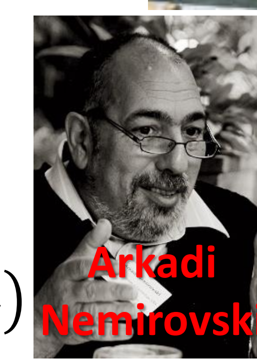
$t^{(k+1)} \leftarrow \mu \cdot t^{(k)}$ (for some parameter $\mu > 1$)

Access to:

- 2nd order oracle for f_0, f_i
- Explicit access to A, b
- Strictly feasible point $x^{(0)}$



- Assumptions:
  - f_0 convex and self-concordant
  - f_i convex quadratic (or linear)
  - $x^{(0)}$ strictly feasible with $f_i\left(x^{(0)}\right) < \delta$
- Overall #Newton Iterations: $O(\sqrt{m}\,(\log {}^1/_\epsilon + \log\log {}^1/_\delta))$
- Overall runtime: $\approx O\left(\sqrt{m}\left((n+p)^3 + m\ \nabla^2 \text{ evals}\right)\log {}^1/_\epsilon\right)$



Optimizing with Matrix Inequalities

$$\boxed{\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & f_0(x) \\ s.t. & f_i(x) \preccurlyeq 0 \\ & Ax = b \end{array}}$$

$$f_i\colon \mathbb{R}^n \to S^{k_i}$$

- Central path given by solutions to:

$$\boxed{\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & f_0(x) - \frac{1}{t}\sum_i \log\det(-f_i(x)) \\ s.t. & Ax = b \end{array}}$$

$$\boxed{\begin{array}{ll} \min\limits_{x \in \mathbb{R}^n} & f_0(x) - \frac{1}{t}\sum_{i=1}^m \log\det\big(-f_i(x)\big) \\ s.t. & Ax = b \end{array}}$$

$$f_i \colon \mathbb{R}^n \to S^{k_i}$$

Optimum x_t^*, dual opt ν_t^*

$$L_t(x, \nu) =$$
$$f_0(x) - \frac{1}{t}\sum_i \log\det\big(-f_i(x)\big) + \langle \nu, Ax - b \rangle$$

$$0 = \nabla_x L_t(x_t^*, \nu_t^*) =$$

$$\nabla f_0(x_t^*) + \sum_i \underbrace{\overset{k_i \times k_i}{\frac{-1}{t} f_i(x_t^*)^{-1}}} \odot \overset{k_i \times k_i \times n}{\nabla f_i(x_t^*)} + A^\top \nu_t^*$$

Define $\lambda_t^* = \frac{-1}{t} f_i(x_t^*)^{-1} \succ 0$

$$\boxed{\begin{array}{ll} \min\limits_{x \in \mathbb{R}^n} & f_0(x) \\ s.t. & f_i(x) \preccurlyeq 0,\ Ax = b \end{array}}$$

x_t^* is strictly feasible

How suboptimal is x_t^* ?

$$L(x, \lambda, \nu) =$$
$$f_0(x) + \sum_i \langle \lambda_i, f_i(x) \rangle + \langle \nu, Ax - b \rangle$$

$$\nabla_x L(x_t^*, \lambda_t^*, \nu_t^*)$$
$$= \nabla f_0(x_t^*) + \sum_i \lambda_t^* \odot \nabla f_i(x_t^*) + A^\top \nu_t^* = 0$$

$$g(\lambda_t^*, \nu_t^*) = \inf_x L(x, \lambda_t^*, \nu_t^*) = L(x_t^*, \lambda_t^*, \nu_t^*)$$
$$= f_0(x_t^*) - \sum \frac{1}{t} \langle f_i(x_t^*)^{-1}, f_i(x_t^*) \rangle + \cancel{\langle \nu_t^*, Ax_t^* - b \rangle}$$
$$= f_0(x_t^*) - \sum \frac{1}{t} tr\, I_{k_i} = f_0(x_t^*) - \frac{\sum_{i=1}^m k_i}{t}$$

λ_t^*, ν_t^* dual (strictly) feasible with

$$f_0(x_t^*) - g(\lambda_t^*, \nu_t^*) = \frac{\sum_{i=1}^m k_i}{t}$$

Optimizing with Matrix Inequalities

$$\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & f_0(x) \\ s.t. & f_i(x) \preccurlyeq 0 \in S^{k_i} \\ & Ax = b \end{array}$$

$$\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & f_0(x) - \frac{1}{t}\sum_i \log\det(-f_i(x)) \\ s.t. & Ax = b \end{array}$$

- An optimum $x^*(t)$ for the t-barrier problem is $\epsilon = \frac{\sum_{i=1}^m k_i}{t}$ suboptimal for constrained problem

Intuition: with single SDP constraint $F \preccurlyeq 0$, $F(x) = diag\big(f_1(x), \dots, f_m(x)\big)$ can handle m scalar constraints $f_i\colon \mathbb{R} \to \mathbb{R}$, so expect error to scale as $\dim(F)$

- Central Path method:

Init: Feasible $x^{(0)}$ and some $t^{(0)}$

Do: Solve $t^{(k)}$-barrier problem using Newton starting at $x^{(k)}$

$x^{(k+1)} \leftarrow x^*(t^{(k)})$

Stop if $\frac{\sum_{i=1}^m k_i}{t} \le \epsilon$

$t^{(k+1)} \leftarrow \mu \cdot t^{(k)}$ (for some parameter $\mu > 1$)

Central Path Method for SDP

Init: Feasible $x^{(0)}$ and some $t^{(0)}$
Do: Solve $t^{(k)}$-barrier problem using Newton starting at $x^{(k)}$
$x^{(k+1)} \leftarrow x^*(t^{(k)})$
Stop if $\frac{\sum_{i=1}^m k_i}{t} \leq \epsilon$
$t^{(k+1)} \leftarrow \mu \cdot t^{(k)}$ (for some parameter $\mu > 1$)

Access to:
- 2nd order oracle for f_0, f_i
- Explicit access to A, b
- Strictly feasible point $x^{(0)}$

- Assumptions:
  - $f_0: \mathbb{R}^n \to \mathbb{R}$ convex and self-concordant
  - $f_i: \mathbb{R}^n \to S^{k_i}$ convex quadratic (or linear)
  - $x^{(0)}$ strictly feasible with $f_i\left(x^{(0)}\right) \prec \delta I$
- Overall #Newton Iterations: $O\left(\sqrt{\sum_i k_i}\,(\log {}^1\!/_\epsilon + \log\log {}^1\!/_\delta)\right)$
- Overall runtime: $\approx O\left(\sqrt{\sum_i k_i}\left((n+p)^3 + \sum_i k_i^3 + m\,\nabla^2 \text{ evals}\right)\log {}^1\!/_\epsilon\right)$

Optimizing with Generalized Inequalities

$$\boxed{\begin{array}{ll} \min_{x\in\mathbb{R}^n} & f_0(x) \\ s.t. & f_i(x) \preccurlyeq_{K_i} 0 \\ & Ax = b \end{array}}$$

$f_i: \mathbb{R}^n \to \mathbb{R}^{k_i}$, proper cone $K_i \subset \mathbb{R}^{k_i}$

<u>Definition</u>: $\psi: \mathbb{R}^k \to \mathbb{R}$ is a **generalized log of degree $\boldsymbol{\theta}$** for the proper cone $K \subset \mathbb{R}^k$ if:

- $\operatorname{dom} \psi = \operatorname{int} K$ (i.e. $\psi(x) < \infty \Leftrightarrow x \succ_K 0$)
- $-\psi$ is closed, **self concordant**, strictly convex ($\nabla^2 \psi(x) \prec 0$)
- For all $x \succ_K 0$ and $s > 0$, we have $\psi(sx) = \psi(x) + \boldsymbol{\theta} \log s$

For $K = S_+^k$: $\log \det(sX) = \log\left(s^k \det X\right) = \log \det X + k \log s$

For positive orthant, $K = \mathbb{R}_+^k$: $\psi(x) = \sum_i \log x_i$

➔ $\psi(sx) = \sum_i \log sx_i = \sum_i(\log x_i + \log s) = \psi(x) + k \log(s)$

Optimizing with Generalized Inequalities

$$\begin{array}{ll} \min_{x\in\mathbb{R}^n} & f_0(x) \\ s.t. & f_i(x) \preccurlyeq_{K_i} 0 \\ & Ax = b \end{array}$$

$$\begin{array}{ll} \min_{x\in\mathbb{R}^n} & f_0(x) - \frac{1}{t}\sum_i \psi_i(f_i) \\ s.t. & Ax = b \end{array}$$

$f_i\colon \mathbb{R}^n \to \mathbb{R}^{k_i}$, proper cone $K_i \subset \mathbb{R}^{k_i}$

$\psi_i\colon \mathbb{R}^{k_i} \to \mathbb{R}$ is a **generalized log of degree θ_i** for K_i

- An optimum $x^*(t)$ for the t-barrier problem is $\epsilon = \frac{\sum_{i=1}^m \theta_i}{t}$ suboptimal for constrained problem
- Central Path method:

Init: Feasible $x^{(0)}$ and some $t^{(0)}$

Do: Solve $t^{(k)}$-barrier problem using Newton starting at $x^{(k)}$

$x^{(k+1)} \leftarrow x^*(t^{(k)})$

Stop if $\frac{\sum_{i=1}^m \theta_i}{t} \leq \epsilon$

$t^{(k+1)} \leftarrow \mu \cdot t^{(k)}$ (for some parameter $\mu > 1$)

Generalized Central Path Method

Init: Feasible $x^{(0)}$ and some $t^{(0)}$

Do: Solve $t^{(k)}$-barrier problem using Newton starting at $x^{(k)}$

$x^{(k+1)} \leftarrow x^*(t^{(k)})$

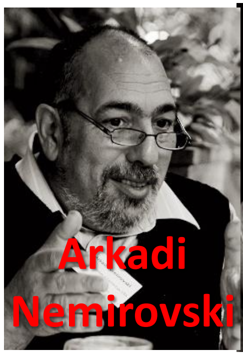
Stop if $\frac{\sum_{i=1}^{m} \theta_i}{t} \leq \epsilon$

$t^{(k+1)} \leftarrow \mu \cdot t^{(k)}$ (for some parameter $\mu > 1$)

Access to:

- 2nd order oracle for f_0, f_i, ψ_i
- Explicit access to A, b
- Strictly feasible point $x^{(0)}$

- Assumptions:
  - ψ_i self-concordant generalized log of degree θ_i for K_i
  - $f_0: \mathbb{R}^n \to \mathbb{R}$ convex and self-concordant
  - $f_i: \mathbb{R}^n \to S^{k_i}$ convex quadratic (or linear)
  - $x^{(0)}$ strictly feasible with $\psi_i\left(-f_i\left(x^{(0)}\right)\right) \leq \log \delta$



How do we find feasible $x^{(0)}$?

- Option 1: use separate procedure

  Phase I: solve feasibility problem to find feasible $x^{(0)}$

  Phase II: solve opt problem starting from $x^{(0)}$

How do we solve a feasibility problem?

Convert to optimization problem (with easy init $x_I^{(0)}$)

(for now, no implicit constraints, $\mathrm{dom}\, f_i = \mathbb{R}^n$, eg LP, QP, SDP, SOCP)

$$\begin{array}{ll} \text{find} & x \in \mathbb{R}^n \\ s.t. & f_i(x) \le 0 \\ & Ax = b \end{array} \Longrightarrow \begin{array}{ll} \min\limits_{x\in\mathbb{R}^n, s\in\mathbb{R}} & s \\ s.t. & f_i(x) \le s \\ & Ax = b \end{array}$$

Initialize to any solution $Ax_I^{(0)} = b$

and $s^{(0)} = \max\limits_i \left(f_i\left(x_I^{(0)}\right)\right) + 1.$

Solve using central path method.

Phase I Method

$$(P)\quad \begin{array}{ll} \text{find} & x \in \mathbb{R}^n \\ s.t. & f_i(x) \le 0 \\ & Ax = b \end{array} \quad \Longrightarrow \quad \begin{array}{ll} \min\limits_{x\in\mathbb{R}^n, s\in\mathbb{R}} & s \\ s.t. & f_i(x) \le s \\ & Ax = b \end{array} \quad (\bar{P})$$

How well do we need to optimize?

- If we find feasible (x, s) for $(\bar{P})$ with $s < 0$ (even if not optimal)
  - ➔ x is strictly feasible for (P), can use as starting point for Phase II
- If we find ϵ-suboptimal (x, s) with $s > \epsilon$
  - ➔ Certificate that (P) is not feasible
- Otherwise (ϵ-suboptimal but $s < \epsilon$)
  - ➔ Not sure, either (P) baerly feasible or barely infeasible

Require δ-strict-feasibility ($\exists_x Ax = b, f_i(x) < \delta$), optimize to $\epsilon = \delta/2$

Phase I+II: Details and Analysis

- To allow bounding overall complexity of Phase I+II, use:

$$\boxed{\begin{array}{lll} \min_{x\in\mathbb{R}^n, s\in\mathbb{R}} & s & \\ s.t. & f_i(x) \le s & Ax = b \\ & \langle g, x\rangle \le 1 & f_0(x) \le M \end{array}}$$

- Assume existence of (unknown) feasible solution with $\|\tilde{x}\| \le R$, $\min_i f_i(\tilde{x}) \le \delta < 0$ and $f_0(\tilde{x}) \le M$
- Pick $x_I^{(0)}$ satisfying $Ax_I^{(0)} = b$ and denote $G = \max_i \left\|\nabla f_i\left(x_I^{(0)}\right)\right\|$ and $F = \max_i f_i(x_I^{(0)}) > 0$
- Use $s_I^{(0)} = mGR + F$ and set $g = -\sum_i \frac{\nabla f_i\left(x_I^{(0)}\right)}{s_I^{(0)} - f_i\left(x_I^{(0)}\right)}$ (this ensures $\langle g, \tilde{x}\rangle \le 1$)

➔ Number of Phase I+II Newton iterations bounded by:

$$N_I = O\left(\sqrt{m}\log\frac{mGR+F}{\delta}\right) \qquad N_{II} = O\left(\sqrt{m}\log\frac{M-p^*}{\epsilon}\right)$$

See Boyd and Vandenberghe Section 11.5.5

With Matrix Inequalities

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f_0(x) \\ s.t. \quad & f_i(x) \preccurlyeq 0 \in S^{k_i} \\ & Ax = b \end{aligned}$$

Phase I Method:

$$\begin{aligned} \text{find} \quad & x \in \mathbb{R}^n \\ s.t. \quad & f_i(x) \preccurlyeq 0 \\ & Ax = b \end{aligned} \quad \Longrightarrow \quad \begin{aligned} \min_{x\in\mathbb{R}^n, s\in\mathbb{R}} \quad & s \\ s.t. \quad & f_i(x) \preccurlyeq sI \\ & Ax = b \end{aligned}$$

Initialize to any solution $Ax_I^{(0)} = b$

and $s_I^{(0)} = \max_i \lambda_{\max}\left(f_i\left(x_I^{(0)}\right)\right) + 1.$

Solve using central path method.

Implicit Constraints?

- Can add as explicit constraints to Phase I method

OR, a an often better alternative to Phase I+II Central Path:

- **Option 2**: Primal-Dual Methods allowing infeasible start

$$\min_{x\in\mathbb{R}^n} \quad f_0(x)$$
$$s.t. \quad f_i(x) \le 0,\ Ax = b$$

$L(x,\lambda,\nu) =$
$f_0(x) + \sum_i \lambda_i f_i(x) + \langle \nu, Ax - b\rangle$

$$\min_{x\in\mathbb{R}^n} \quad f_0(x) - \frac{1}{t}\sum_{i=1}^m \log(-f_i(x))$$
$$s.t. \quad Ax = b$$

$L_t(x,\nu) =$
$f_0(x) - \frac{1}{t}\sum_i \log(-f_i(x)) + \langle \nu, Ax - b\rangle$

Newton iteration of log-barrier method:

1. Use (C) to eliminate $\lambda_i = \frac{-1}{tf_i(x)}$, substitute to get problem in x, ν.

➔ (D) becomes $0 = \nabla_x L\left(x, \frac{-1}{tf_i(x)}, \nu\right) = \nabla_x L_t(x,\nu)$

2. Linearize (D) about $x = x^{(k)} + \Delta x$
3. Solve (P)+($\widetilde{\text{D}}$) for $\Delta x, \nu$, take step in direction Δx (ensuring progress, and (f)+(λ) remain valid)

Relaxed KKT

(f) $f_i(x) \le 0$

(λ) $\lambda_i \ge 0$

(P) $Ax = b$

(D) $\nabla_x L(x,\lambda,\nu) = 0$

(C) $\lambda_i \cdot f_i(x) = -{}^1/_t$

Iteration of Primal/Dual method: (work on $x^{(k)}, \lambda^{(k)}, \nu^{(k)}$)

1. Linearize (D)+(C) about $x = x^{(k)} + \Delta x$ and $\lambda = \lambda^{(k)} + \Delta x$
2. Solve (P)+($\widetilde{\text{D}}$)+($\tilde{\text{C}}$) for $x = x^{(k)} + \Delta x, \lambda = \lambda^{(k)} + \Delta, \nu = \nu^{(k)} + \Delta\,\nu$ take step in direction $\Delta x, \Delta\lambda\,, \Delta\nu$ (ensuring (f)+(λ) remain valid)