Ch. 11. Interior-point methods.

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \le 0, \quad i=1,\dots,m \qquad (1) \\ & Ax = b. \end{aligned}$$

- $f_i(x)$ convex, twice continuously differentiable.
- $A \in \mathbb{R}^{q \times n}$. $\operatorname{rank}(A) = q$.
- We assume x^* exists. and $f_0(x^*) = p^*$ is finite.
- assume that the problem is strictly feasible.

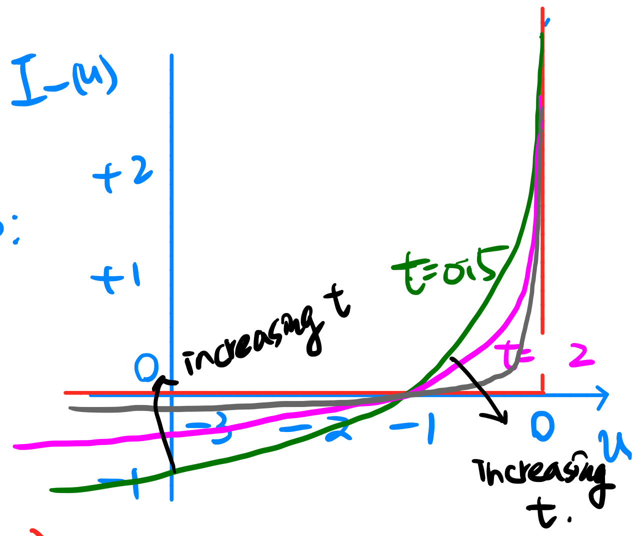
  i.e $\exists\ \hat{x} \in \operatorname{dom} f_0$, $f_i(\hat{x}) < 0$. $i=1,\dots,m$

  $A\hat{x} = b.$

Logarithmic barrier

Define $I_-(u) = \begin{cases} 0, & \text{if } u \le 0 \\ \infty, & \text{otherwise.} \end{cases}$

(indication function of $\mathbb{R}_-$).



Reformulate (1) as follows:

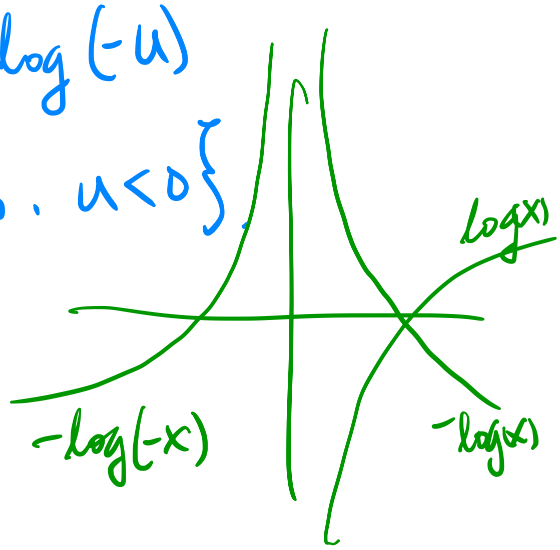
$$\text{minimize } f_0(x) + \sum_{i=1}^{m} I_-(f_i(x))$$

s.t. $Ax = b$.

If x strictly feasible. $f_i(x) < 0$

- For $t > 0$, $\hat{I}_-(u) = -\frac{1}{t}\log(-u)$

with $\operatorname{dom} \hat{I}_- = \{u \in \mathbb{R}.\ u < 0\}$.



- approximation via logarithmic barrier

$$\text{minimize} \quad f_0(x) - \frac{1}{t}\sum_{i=1}^{m} \log(-f_i(x))$$

$$\text{s.t.} \quad Ax = b.$$

- an equality constrained function.
- As t increases, the approximation becomes more accurate.
- Logarithmic barrier function (log barrier)

$$\phi(x) = -\sum_{i=1}^{m} \log(-f_i(x)).$$

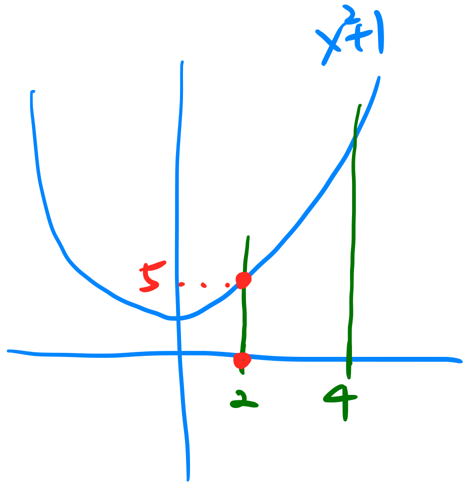
$$\text{dom}\,\phi = \{x : f_1(x) < 0, \ldots, f_m(x) < 0\}.$$

- $\phi(x)$ convex. by the composition rule.
- $\phi(x)$ is twice continuously differentiable.

$$\begin{cases} \nabla\phi(x) = \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla f_i(x). \\ \nabla^2\phi(x) = \sum_{i=1}^{m} \frac{1}{f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T + \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla^2 f_i(x). \end{cases}$$

Ex. minimize x^2+1

s.t. $2 \le x \le 4$.



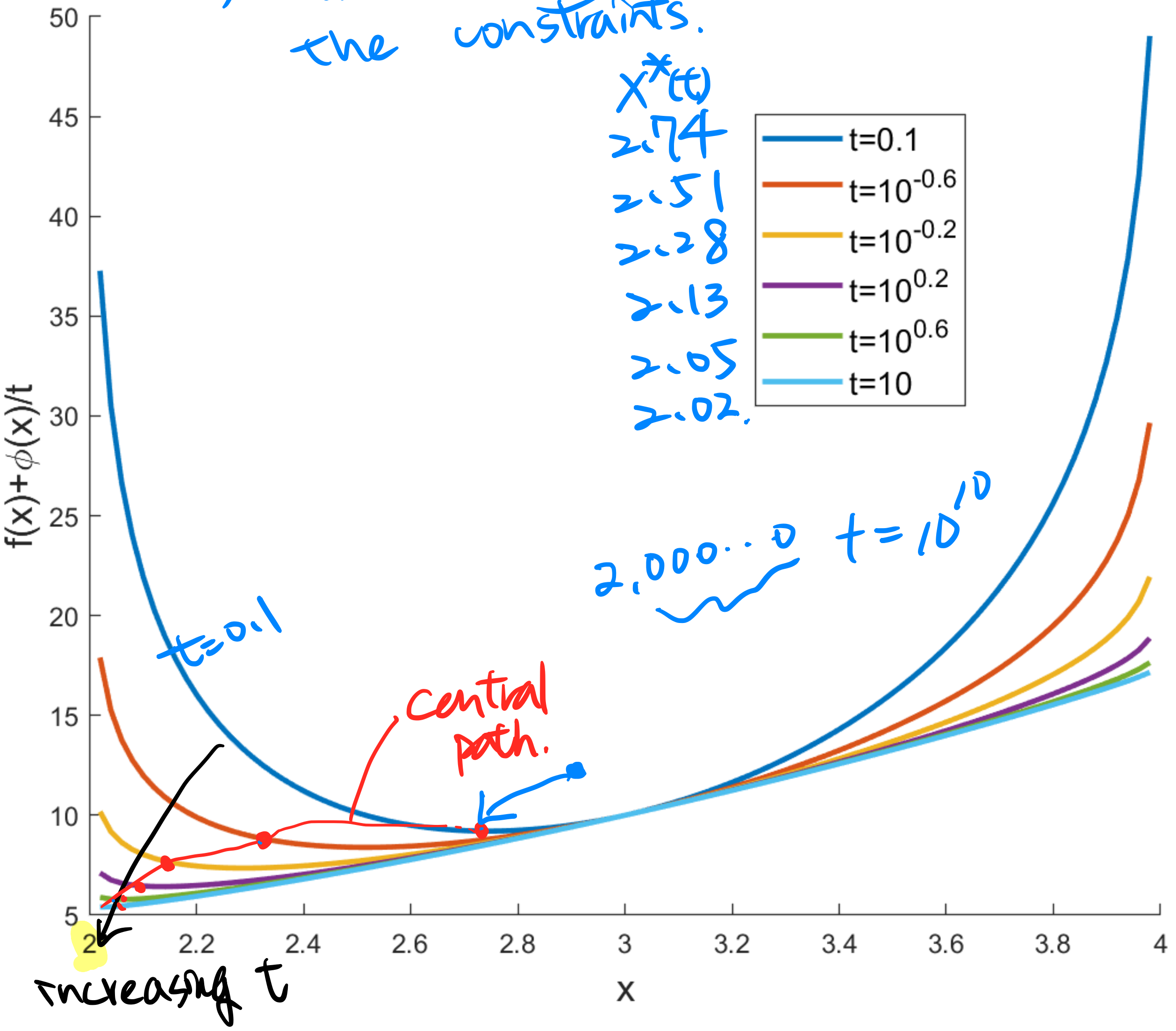
Reformulate the problem as

$$\text{minimize} \quad \underbrace{x^2+1-\frac{1}{t}\log(4-x)-\frac{1}{t}\log(x-2)}$$

$(t>0)$.

$$0=\frac{\partial}{\partial x}(\quad)=2tx+\frac{1}{4-x}-\frac{1}{x-2}$$

$\Rightarrow$ three roots. but only one fits the constraints.



Central path.

for $t > 0$, define $x^*(t)$ as the solution of

$$\begin{aligned} \text{minimize} \quad & t f_0(x) + \phi(x). \\ \text{s.t.} \quad & Ax = b. \end{aligned}$$

$L(x, \nu) = t f_0(x) + \phi(x) + \nu^T(Ax - b)$

(From now on, assume $x^*(t)$ exists and is unique for $t > 0$.)

— the central path associated with (1) is $\{x^*(t) : t > 0\}$.

— $x^*(t)$ is strictly feasible.

$Ax^*(t) = b$, $f_i(x^*(t)) < 0$, $i = 1, \ldots, m$.

if and only if there exists $\hat{\nu}$

$$0 = t \nabla f_0(x) + \nabla \phi(x) + A^T \hat{\nu}$$

Δx.

at $x = x^*(t)$.

$Ax^*(t) = b$.

(centrality condition).

Example central path for an LP.

$$\text{minimize} \quad c^T x$$
$$\text{s.t.} \quad a_i^T x \le b_i \;, \; i=1,\dots,m.$$

— log barrier $\phi(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x)$,

$$\operatorname{dom} \phi = \{ x : Ax < b \}$$

where $A = \begin{bmatrix} — a_1^T — \\ — a_2^T — \\ \vdots \\ — a_m^T — \end{bmatrix}$

— The centrality condition is

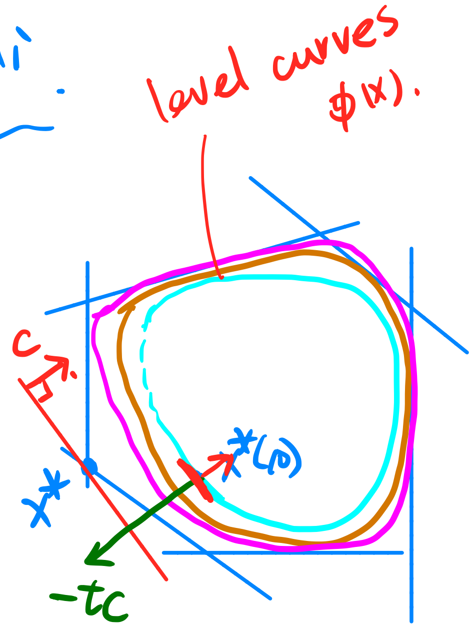
$$L(x) = c^T x + \frac{1}{t} \phi(x).$$

$$0 = \frac{\partial L(x)}{\partial x_j} = c_j + \frac{1}{t} \sum_{i=1}^{m} \frac{a_{ij}}{b_i - a_i^T x}.$$

$$\Rightarrow 0 = tc + \underbrace{\sum_{i=1}^{m} \frac{1}{b_i - a_i^T x} a_i}_{\nabla \phi(x)}.$$

$$\nabla \Phi(x^*(t)) = -tc \;, \quad t > 0$$

which is parallel to $-c$.



— The hyperplane $\boxed{c^T x = c^T x^*(t)}$ is tangent to the level curve of $\phi(x)$ through $x^*(t)$.

$$c^T (x - x^*(t)) = 0.$$

Dual points on central path.

$x^*(t)$ if there ν^* such that

$$\begin{cases} \underline{t}\nabla f_0(x^*(t)) + \sum_{i=1}^m \boxed{\dfrac{1}{-tf_i(x^*(t))}} \nabla f_i(x^*(t)) + A^T \boxed{\dfrac{\nu^*}{t}} = 0 \\ \quad t \\ A x^*(t) = 0. \end{cases}$$

Therefore $x^*(t)$ minimizes the Lagragian

of (1).

$$L(x, \lambda^*(t), \nu^*(t)) = f_0(x) + \sum_{i=1}^m \lambda_i^*(t) f_i(x) + \nu^*(t)^T (Ax-b)$$

where $\begin{cases} \lambda_i^*(t) \triangleq \dfrac{1}{-tf_i(x^*(t))} \\ \nu^*(t) = \dfrac{\nu^*}{t}. \end{cases}$

This means, that $(\lambda_i^*(t), \nu^*(t))$ is a dual feasible pair.

$$g(\lambda^*(t), \nu^*(t)) = f_0(x^*(t)) + \sum_{i=1}^m \underbrace{\lambda_i^*(t) f_i(x^*(t))}_{\frac{1}{-t}} + \nu^*(t)^T \underbrace{(Ax^*(t)-b)}_{=0}$$

$$= f_0(x^*(t)) - \boxed{\frac{m}{t}}$$

duality gap.

$$p^* \geq g(\lambda^*(t), \nu^*(t)) = f_0(x^*(t)) - \frac{m}{t}$$

$$\Rightarrow \quad f_0(x^*(t)) - p^* \leq \boxed{\frac{m}{t}.} \to 0$$

as $\to \infty$.

Barrier method

given strictly feasible x, $t := t^{(0)} > 0$, $\mu > 1$, tolerance $\epsilon > 0$.

repeat

1. Centering step.
   Compute $x^*(t)$ by minimizing $t f_0 + \phi$ s.t. $Ax = b$, starting at x.
2. Update $x := x^*(t)$.
3. Stopping criterion. quit if $\frac{m}{t} \le \epsilon$
4. Increase $t := \mu t$.

- choices of μ: usually $10 \sim 20$.
- several heuristics for choice of $t^{(0)}$.
- centering usually done using Newton's method, starting at current x.

The Newton step Δx_{nt}

$$\begin{bmatrix} t\nabla^2 f_0(x) + \nabla^2 \phi(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{nt} \\ \nu_{nt} \end{bmatrix} = - \begin{bmatrix} t\nabla f_0(x) + \nabla\phi(x) \\ 0 \end{bmatrix}$$

Convergence analysis.

centering problem is to

$$\begin{aligned} \text{minimize} \quad & t f_0(x) + \phi(x) \\ \text{s.t.} \quad & Ax = b. \end{aligned}$$

$t^{(0)}, \quad \mu t^{(0)}, \quad \mu^2 t^{(0)}, \ \cdots, \ \cdots, \mu^k t^{(0)}, \cdots$

The duality gap is $\dfrac{m}{t} = \dfrac{m}{\mu^k t^{(0)}}$

$$\frac{m}{\mu^k t^{(0)}} < \epsilon \Rightarrow \frac{m}{t^{(0)}\epsilon} < \mu^k$$

$$\Rightarrow \log \frac{m}{t^{(0)}\epsilon} < k \log \mu.$$

$$\Rightarrow k \geq \left\lceil \frac{\log m/t^{(0)}\epsilon}{\log \mu} \right\rceil$$

Therefore the desired accuracy ϵ is achieved after exactly $\left\lceil \dfrac{\log m/t^{(0)}\epsilon}{\log \mu} \right\rceil$ centering steps, plus the initial centering step.