12. Interior-point methods

- inequality constrained minimization
- logarithmic barrier function and central path
- barrier method
- feasibility and phase I methods
- generalized inequalities
Inequality constrained minimization

$$\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& \quad Ax = b
\end{align*}$$

(1)

- $f_i$ convex, twice continuously differentiable
- $A \in \mathbb{R}^{p \times n}$ with $\text{rank} \, A = p$
- we assume $p^*$ is finite and attained
- we assume problem is strictly feasible: there exists $\tilde{x}$ with

$$\tilde{x} \in \text{dom} \, f_0, \quad f_i(\tilde{x}) < 0, \quad i = 1, \ldots, m, \quad A\tilde{x} = b$$

(Slater's condition)

hence, strong duality holds and dual optimum is attained
Examples

- LP, QP, QCQP, GP

- entropy maximization with linear inequality constraints

  \[
  \begin{align*}
  \text{minimize} & \quad \sum_{i=1}^{n} x_i \log x_i \\
  \text{subject to} & \quad Fx \preceq g \\
  & \quad Ax = b
  \end{align*}
  \]

  with \( \text{dom } f_0 = \mathbb{R}^n_{++} \)

- differentiability may require reformulating the problem, \( e.g., \) piecewise-linear minimization or \( \ell_\infty \)-norm approximation via LP

- SDPs and SOCPs are better handled as problems with generalized inequalities (see later)
Logarithmic barrier

reformulation of (1) via indicator function:

\[
\begin{align*}
\text{minimize} & \quad f_0(x) + \sum_{i=1}^{m} I_-(f_i(x)) \\
\text{subject to} & \quad Ax = b
\end{align*}
\]

where \( I_- (u) = 0 \) if \( u \leq 0 \), \( I_- (u) = \infty \) otherwise (indicator function of \( \mathbb{R}_- \))

approximation via logarithmic barrier

\[
\begin{align*}
\text{minimize} & \quad f_0(x) - \frac{1}{t} \sum_{i=1}^{m} \log(-f_i(x)) \\
\text{subject to} & \quad Ax = b
\end{align*}
\]

• an equality constrained problem

• for \( t > 0 \), \( -(1/t) \log(-u) \) is a smooth approximation of \( I_- \)

• approximation improves as \( t \to \infty \)
logarithmic barrier function

\[ \phi(x) = -\sum_{i=1}^{m} \log(-f_i(x)), \quad \text{dom} \phi = \{ x \mid f_1(x) < 0, \ldots, f_m(x) < 0 \} \]

(Slater’s condition)

- convex (follows from composition rules)
- twice continuously differentiable, with derivatives

\[ \nabla \phi(x) = \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla f_i(x) \]

\[ \nabla^2 \phi(x) = \sum_{i=1}^{m} \frac{1}{f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T + \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla^2 f_i(x) \]

(Useful for KKT analysis and Newton’s method)
Central path

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

$$
\begin{align*}
& \text{minimize} \quad tf_0(x) + \phi(x) \\
& \text{subject to} \quad Ax = b
\end{align*}
$$

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

- central path is $\{x^*(t) \mid t > 0\}$

**example:** central path for an LP

$$
\begin{align*}
& \text{minimize} \quad c^T x \\
& \text{subject to} \quad a_i^T x \leq b_i, \quad i = 1, \ldots, 6
\end{align*}
$$

hyperplane $c^T x = c^T x^*(t)$ is tangent to level curve of $\phi$ through $x^*(t)$

Also, as $t$ increases, we obtain $x^*(t)$ approaches the optimal of the original problem.
Dual points on central path

\[ x = x^*(t) \] if there exists a \( w \) such that

\[
t \nabla f_0(x) + \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla f_i(x) + A^T w = 0,
\]

\[ Ax = b \quad \text{(Primal feasibility)} \]

• therefore, \( x^*(t) \) minimizes the Lagrangian

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

where we define \( \lambda_i^*(t) = 1/(-t f_i(x^*(t))) \) and \( \nu^*(t) = w/t \)

> 0 since \( t > 0 \) and \( f_i(x^*(t)) < 0 \)

• this confirms the intuitive idea that \( f_0(x^*(t)) \rightarrow p^* \) if \( t \rightarrow \infty \):

\[
p^* \geq g(\lambda^*(t), \nu^*(t)) \quad \text{... for any (\lambda, \nu) so we can plug (\lambda^*(t), \nu^*(t))}
\]

\[
= L(x^*(t), \lambda^*(t), \nu^*(t))
\]

\[
= f_0(x^*(t)) + \sum_i \lambda_i^*(t) f_i(x^*(t)) + \nu^*(t)(A x^*(t) - b)
\]

\[
= f_0(x^*(t)) + \sum_i f_i(x^*(t)) / (-t f_i(x^*(t))) \quad \text{... since } A x^*(t) = b
\]

\[
= f_0(x^*(t)) - m/t \quad \text{... m terms}
\]

As \( t \rightarrow \infty \), \( m/t \rightarrow 0 \) and then \( p^* = f_0(x^*(t)) \).
Interpretation via KKT conditions

\[ x = x^*(t), \lambda = \lambda^*(t), \nu = \nu^*(t) \text{ satisfy} \]

1. primal constraints: \( f_i(x) \leq 0, i = 1, \ldots, m, \ Ax = b \)
2. dual constraints: \( \lambda \geq 0 \)
3. approximate complementary slackness: \(-\lambda_i f_i(x) = \frac{1}{t}, i = 1, \ldots, m\)
4. gradient of Lagrangian with respect to \( x \) vanishes:

\[
\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + A^T \nu = 0
\]

difference with KKT is that condition 3 replaces \( \lambda_i f_i(x) = 0 \)

We said before: \( \lambda_i(t) = \frac{1}{t f_i(x)} \)

Recall original problem:
\[
\min f_0(x)
\text{s.t.} \ f_i(x) \leq 0, \ i = 1 \ldots m
\Ax = b
\]
Barrier method

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

repeat

1. Centering step. Compute \( x^*(t) \) by minimizing \( tf_0 + \phi, \) subject to \( Ax = b. \)
2. Update. \( x := x^*(t). \)
3. Stopping criterion. \textbf{quit} if \( m/t < \epsilon. \)
4. Increase \( t. \) \( t := \mu t. \)

- terminates with \( f_0(x) - p^* \leq \epsilon \) (stopping criterion follows from \( f_0(x^*(t)) - p^* \leq m/t \))

- centering usually done using Newton’s method, starting at current \( x \)

The gradient at the current \( x \) is \( d = t dfo(x) + d\phi(x) \)
The Hessian at the current \( x \) is \( H = t d^2fo(x) + d^2\phi(x) \)

\[
[H \ A'] [\Delta x] = [-d] \\
[A \ 0] [v] = [0]
\]
Feasibility and phase I methods

**feasibility problem:** find $x$ such that

$$f_i(x) \leq 0, \quad i = 1, \ldots, m, \quad Ax = b$$

(2)

**phase I:** computes strictly feasible starting point for barrier method

**basic phase I method**

minimize (over $x, s$) $s$
subject to $f_i(x) \leq s, \quad i = 1, \ldots, m$
$Ax = b$

(3)

• if $x, s$ feasible, with $s < 0$, then $x$ is strictly feasible for (2)
• if optimal value $\bar{p}^*$ of (3) is positive, then problem (2) is infeasible $(s>0)$
• if $\bar{p}^* = 0$ and attained, then problem (2) is feasible (but not strictly);
  if $\bar{p}^* = 0$ and not attained, then problem (2) is infeasible
Generalized inequalities

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \preceq K_i 0, \quad i = 1, \ldots, m \\
& \quad Ax = b
\end{align*}
\]

- \( f_0 \) convex, \( f_i : \mathbb{R}^n \to \mathbb{R}^{k_i} \), \( i = 1, \ldots, m \), convex with respect to proper cones \( K_i \in \mathbb{R}^{k_i} \)
- \( f_i \) twice continuously differentiable
- \( A \in \mathbb{R}^{p \times n} \) with rank \( A = p \)
- we assume \( p^* \) is finite and attained
- we assume problem is strictly feasible; hence strong duality holds and dual optimum is attained

examples of greatest interest: SOCP, SDP
Generalized logarithm for proper cone

\( \psi : \mathbb{R}^q \to \mathbb{R} \) is generalized logarithm for proper cone \( K \subseteq \mathbb{R}^q \) if:

- \( \text{dom} \psi = \text{int} K \) and \( \nabla^2 \psi(y) \prec 0 \) for \( y \succ_K 0 \)
- \( \psi(sy) = \psi(y) + \theta \log s \) for \( y \succ_K 0, \, s > 0 \) \( (\theta > 0 \text{ is the degree of } \psi) \)

\[ \text{Examples} \]

- Nonnegative orthant \( K = \mathbb{R}_+^n \): \( \psi(y) = \sum_{i=1}^{n} \log y_i \), with degree \( \theta = n \)
- Positive semidefinite cone \( K = S_+^n \):

  \[ \psi(Y) = \log \det Y \quad (\theta = n) \]

- Second-order cone \( K = \{ y \in \mathbb{R}^{n+1} | \left( y_1^2 + \cdots + y_n^2 \right)^{1/2} \leq y_{n+1} \} \):

  \[ \psi(y) = \log(y_{n+1}^2 - y_1^2 - \cdots - y_n^2) \quad (\theta = 2) \]

Take \( K = \{ z \in \mathbb{R} | z \geq 0 \} \): \( \psi(z) = \log z \)

For \( y > 0, \, s > 0 \):

\[ \psi(sy) = \psi(y) + \theta \log s, \text{ where } \theta = 1 \]
Recall proper cones (2-21):
\[
z \geq_{K^*} 0 \text{ if and only if } y'z \geq 0 \text{ for all } y \geq_{K^*} 0
\]

for \( y \succ K^0 \),
\[
\nabla \psi(y) \succeq_{K^*} 0, \quad y^T \nabla \psi(y) = \theta
\]

\[
\nabla \psi(Y) = Y^{-1}, \quad \text{tr}(Y \nabla \psi(Y)) = n
\]

\[
\nabla \psi(y) = \frac{2}{y_{n+1}^2 - y_1^2 - \cdots - y_n^2}
\begin{bmatrix}
-y_1 \\
\vdots \\
-y_n \\
y_{n+1}
\end{bmatrix}, \quad y^T \nabla \psi(y) = 2
\]

\[
\text{take } s=1: \quad y' d\psi(y) = \theta
\]

Recall:
\[
\psi(s\ y) = \psi(y) + \theta \log s
\]

from left:
\[
d\psi(s\ y)/ds = y' d\psi(s\ y)
\]

from right:
\[
d\psi(s\ y)/ds = \theta/s
\]

thus:
\[
y' d\psi(s\ y) = \theta/s
\]

Indeed \( y' d\psi(y) = \theta > 0 \)

• nonnegative orthant \( \mathbb{R}^n_+ \):
\[
\psi(y) = \sum_{i=1}^n \log y_i
\]

\[
\nabla \psi(y) = (1/y_1, \ldots, 1/y_n), \quad y^T \nabla \psi(y) = n
\]

\[
\text{tr}(Y' d\psi(Y)) = \text{tr} \ 1
\]

• positive semidefinite cone \( \mathbb{S}^n_+ \):
\[
\psi(Y) = \log \det Y
\]

\[
\nabla \psi(Y) = Y^{-1}
\]

\[
\text{tr}(Y \nabla \psi(Y)) = n
\]

• second-order cone \( K = \{ y \in \mathbb{R}^{n+1} | (y_1^2 + \cdots + y_n^2)^{1/2} \leq y_{n+1} \} \):

\[
\text{Interior-point methods}
\]

12–25
Logarithmic barrier and central path

**logarithmic barrier** for $f_1(x) \preceq_{K_1} 0, \ldots, f_m(x) \preceq_{K_m} 0$:

$$
\phi(x) = -\sum_{i=1}^{m} \psi_i(-f_i(x)), \quad \text{dom} \phi = \{ x \mid f_i(x) \prec_{K_i} 0, \ i = 1, \ldots, m \}
$$

- $\psi_i$ is generalized logarithm for $K_i$, with degree $\theta_i$
- $\phi$ is convex, twice continuously differentiable

**central path**: $\{x^*(t) \mid t > 0\}$ where $x^*(t)$ solves

$$
\begin{align*}
\text{minimize} & \quad tf_0(x) + \phi(x) \\
\text{subject to} & \quad Ax = b
\end{align*}
$$
Dual points on central path

\[ x = x^*(t) \] if there exists \( w \in \mathbb{R}^p \),

\[
t \nabla f_0(x) + \sum_{i=1}^{m} Df_i(x)^T \nabla \psi_i(-f_i(x)) + A^T w = 0
\]

\( Df_i(x) \in \mathbb{R}^{k_i \times n} \) is derivative matrix of \( f_i : \mathbb{R}^n \to \mathbb{R}^{\{k_i\}} \)

- therefore, \( x^*(t) \) minimizes Lagrangian \( L(x, \lambda^*(t), \nu^*(t)) \), where

\[
\lambda_i^*(t) = \frac{1}{t} \nabla \psi_i(-f_i(x^*(t))), \quad \nu^*(t) = \frac{w}{t}
\]

- from properties of \( \psi_i \): \( \lambda_i^*(t) \succ_{K_i^*} 0 \), with duality gap

\[
f_0(x^*(t)) - g(\lambda^*(t), \nu^*(t)) = \frac{1}{t} \sum_{i=1}^{m} \theta_i
\]

\( p^* \geq g(\lambda^*(t), \nu^*(t)) \) ... for any \((\lambda, \nu)\) so we can plug \((\lambda^*(t), \nu^*(t))\)

\[
P^* = L(x^*(t), \lambda^*(t), \nu^*(t)) = f_0(x^*(t)) + \sum_i \lambda_i^* \nu_i^* f_i(x^*(t)) + \nu^*(t)(A x^*(t) - b)
\]

\[
= f_0(x^*(t)) - \frac{1}{t} \sum_i y_i \psi_i(y_i) \quad \text{... since } A x^*(t) = b \text{, and letting } y_i = -f_i(x^*(t))
\]

\[
= f_0(x^*(t)) - \frac{1}{t} \sum_i \theta_i \quad \text{... since } y_i \psi_i(y_i) = \theta_i
\]

Rest of the text is not visible.
Barrier method

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

repeat
1. Centering step. Compute $x^*(t)$ by minimizing $tf_0 + \phi$, subject to $Ax = b$.
2. Update. $x := x^*(t)$.
3. Stopping criterion. quit if $(\sum_i \theta_i)/t < \epsilon$.
4. Increase $t$. $t := \mu t$.

- only difference is duality gap $m/t$ on central path is replaced by $\sum_i \theta_i/t$