12. Interior-point methods

- inequality constrained minimization
- logarithmic barrier function and central path
- barrier method
- feasibility and phase I methods
- complexity analysis via self-concordance
- generalized inequalities

Inequality constrained minimization

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

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

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

hence, strong duality holds and dual optimum is attained

Examples

- LP, QP, QCQP, GP

- entropy maximization with linear inequality constraints

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n x_i \log x_i \\ \text{subject to} & Fx \preceq g \\ & Ax = b \end{array}$$

  with $\mathbf{dom}\, f_0 = \mathbf{R}^n_{++}$

- differentiability may require reformulating the problem, $e.g.$, piecewise-linear minimization or ℓ_∞-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:

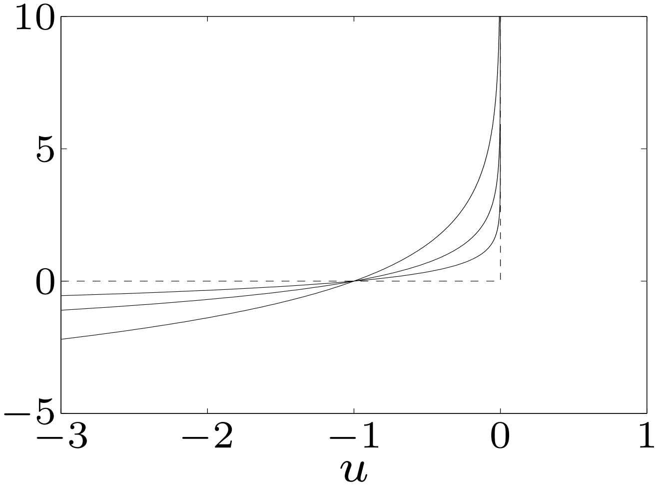
$$\begin{array}{ll} \text{minimize} & f_0(x) + \sum_{i=1}^m I_-(f_i(x)) \\ \text{subject to} & Ax = b \end{array}$$

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

approximation via logarithmic barrier

$$\begin{array}{ll} \text{minimize} & f_0(x) - (1/t) \sum_{i=1}^m \log(-f_i(x)) \\ \text{subject to} & Ax = b \end{array}$$

- 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

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

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

$$\begin{array}{rcl} \nabla\phi(x) & = & \displaystyle\sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla f_i(x) \\ \nabla^2\phi(x) & = & \displaystyle\sum_{i=1}^{m} \frac{1}{f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T + \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla^2 f_i(x) \end{array}$$

Central path

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

$$\begin{array}{ll} \text{minimize} & t f_0(x) + \phi(x) \\ \text{subject to} & Ax = b \end{array}$$

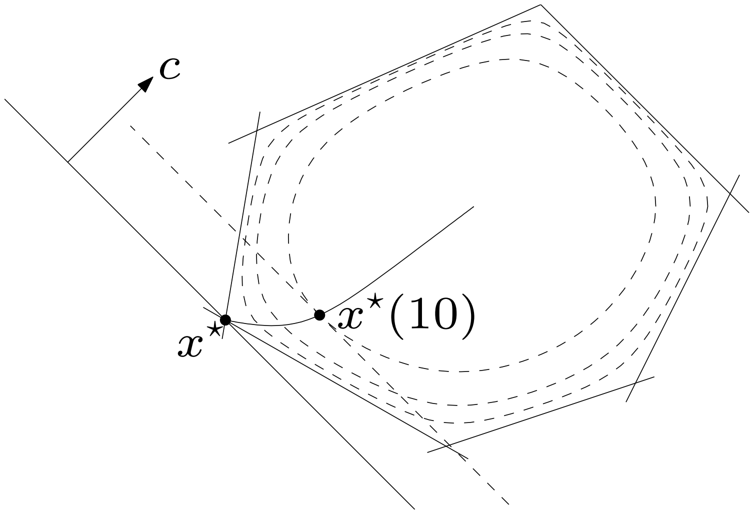
(for now, assume $x^\star(t)$ exists and is unique for each $t > 0$)

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

example: central path for an LP

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

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



Dual points on central path

$x = x^\star(t)$ if there exists a w such that

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

- therefore, $x^\star(t)$ minimizes the Lagrangian

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

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

- this confirms the intuitive idea that $f_0(x^\star(t)) \to p^\star$ if $t \to \infty$:

$$\begin{aligned} p^\star &\geq g(\lambda^\star(t), \nu^\star(t)) \\ &= L(x^\star(t), \lambda^\star(t), \nu^\star(t)) \\ &= f_0(x^\star(t)) - m/t \end{aligned}$$

Interpretation via KKT conditions

$x = x^\star(t)$, $\lambda = \lambda^\star(t)$, $\nu = \nu^\star(t)$ satisfy

1. primal constraints: $f_i(x) \leq 0$, $i = 1, \ldots, m$, $Ax = b$
2. dual constraints: $\lambda \succeq 0$
3. approximate complementary slackness: $-\lambda_i f_i(x) = 1/t$, $i = 1, \ldots, m$
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 $\lambda_i f_i(x) = 0$