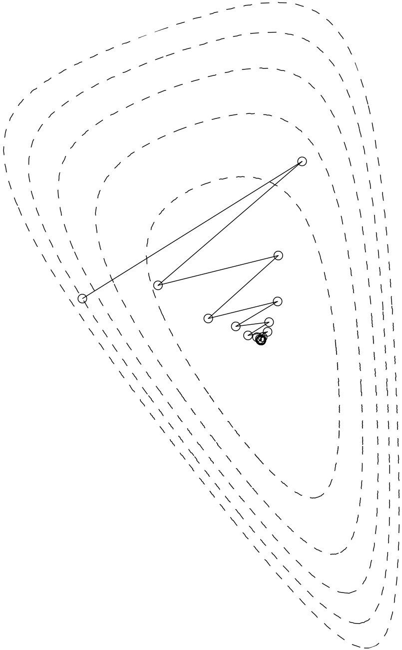


Optimization Algorithms

Log Barrier Method

Log barrier, central path, relaxed KKT

Marc Toussaint
Technical University of Berlin
Winter 2024/25

Barriers & Penalties

- The general approach is to define unconstrained problems that help tackling the constrained problem
  - We typically add penalties or barriers to the cost function
  - Iteratively solving the unconstrained problem converges to the solution of the constrained problem

- A **barrier** is ∞ for $g(x) > 0$

- A **penalty** is zero for $g(x) \leq 0$ and increases with $g(x) > 0$

Unconstrained "inner" problems to tackle constraints

- General approach is to *augment* $f(x)$ with some terms to define an "inner" problem:

$$B(x,\mu) = f(x) - \mu \sum_i \log(-g_i(x)) \qquad \text{(log barrier)}$$

$$S(x,\mu,\nu) = f(x) + \mu \sum_i [g_i(x) > 0]\ g_i(x)^2 + \nu \sum_j h_j(x)^2 \qquad \text{(sqr. penalty)}$$

$$L(x,\lambda,\kappa) = f(x) + \sum_i \lambda_i g_i(x) + \sum_j \kappa_j h_j(x) \qquad \text{(Lagrangian)}$$

$$\begin{aligned} A(x,\lambda,\kappa,\mu,\nu) = f(x) &+ \sum_i \lambda_i g_i(x) + \sum_j \kappa_j h_j(x) + \\ &+ \mu \sum_i [g_i(x) > 0]\ g_i(x)^2 + \nu \sum_j h_j(x)^2 \qquad \text{(Aug. Lag.)} \end{aligned}$$

Log barrier (Interior Point) method

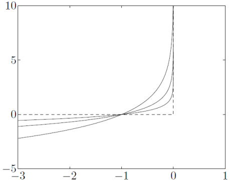
- To solve the original problem

$$\min_x \; f(x) \;\; \text{s.t.} \;\; g(x) \le 0$$

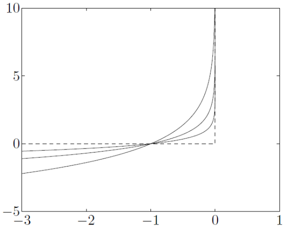
we define the unconstrained *inner* problem

$$\min_x B(x,\mu)\;, \qquad B(x,\mu) = f(x) - \mu \sum_i \log(-g_i(x))$$

with **log barrier** function $-\mu \log(-g)$:



Log barrier



- For $\mu \to 0$, the log barrier $-\mu \log(-g)$ converges to $\infty[g > 0]$

Notation: [*boolean expression*] $\in \{0, 1\}$

- The neg barrier gradient $\nabla \log(-g) = \frac{\nabla g}{g}$ pushes away from the constraint
- Eventually we want to have a very small μ
  - But choosing small μ from the start makes the barrier "non-smooth"

$\to$ gradually *decrease* μ

Log barrier method

Input: initial $x \in \mathbb{R}^n$, functions $f(x), g(x)$, tolerance θ, parameters (defaults: $\varrho_\mu^- = 0.5, \mu_0 = 1$)
Output: x

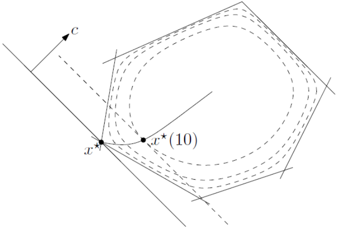
1: initialize $\mu = \mu_0$
2: **repeat**
3: **centering: solve unconstrained problem** $x \leftarrow \text{argmin}_x \; B(x, \mu)$
4: decrease $\mu \leftarrow \varrho_\mu^- \mu$
5: **until** $|\Delta x| < \theta$ repeatedly

- Note: See Boyd & Vandenberghe for alternative stopping criteria based on f precision (duality gap) and better choice of initial μ (which is called t there).

- For reference: $\nabla_x B(x, \mu) = \nabla f(x) - \mu \sum_i \frac{1}{g_i(x)} \; \nabla g_i(x) \;, \quad \nabla_x^2 B(x, \mu) \approx \nabla^2 f(x) + \mu \sum_i \frac{1}{g_i(x)^2} \; \nabla g_i(x) \nabla g_i(x)^\top$

Central Path

- Every μ defines a different $x^*(\mu) = \text{argmin}_x \; B(x, \mu)$



- Varying μ creates the **Central path** with points $x^*(\mu)$, gradually approaching the optimum for $\mu \to 0$ from the **interior**
- This is an **Interior Point Method**: all iterates will always fulfill $g_i(x) < 0$

Comments

- We always have to initialize log barrier with an interior point
- Equality constraints need to be handled separately (e.g. AugLag)

Log barrier $\leftrightarrow$ relaxed KKT

- Let's look at the gradients at the optimum $x^*(\mu) = \text{argmin}_x B(x, \mu)$

$$\nabla f(x) - \sum_i \frac{\mu}{g_i(x)} \nabla g_i(x) = 0$$

$$\Leftrightarrow \qquad \nabla f(x) + \sum_i \lambda_i \nabla g_i(x) = 0 \;, \quad \lambda_i g_i(x) = -\mu$$

  where we defined(!) $\lambda_i = -\mu / g_i(x)$, which guarantees $\lambda_i \geq 0$ as long as we are in the interior ($g_i \leq 0$)

- These are called **modified (=relaxed) KKT conditions** with **relaxed complementarity**
  - Inequalities also push in the interior
  - For $\mu \to 0$ we converge to the exact KKT conditions

Log barrier $\leftrightarrow$ relaxed KKT

- *Centering (solving the unconstrained problem) in the log barrier method is equivalent to solving the relaxed KKT conditions!*

(Reference for later: μ can be interpreted as upper bound on the sub-optimality).

Log barrier $\leftrightarrow$ relaxed KKT

- *Centering (solving the unconstrained problem) in the log barrier method is equivalent to solving the relaxed KKT conditions!*

  (Reference for later: μ can be interpreted as upper bound on the sub-optimality).

- Nice about the log barrier method:
  - It "smoothes out" the combinatorics of constraint activity, smoothly approaching the optimum from the interior
  - This smoothing mathematically relates to relaxing the complementarity condition: constraints can push also in the interior, μ relates to the degree of smoothing/relaxation

- Demo...