

Barrier Method

Javier Peña

Convex Optimization 10-725/36-725

Last time: Newton's method

For root-finding

$$F(x) = 0 \rightsquigarrow x^+ = x - F'(x)^{-1}F(x)$$

For optimization

$$\min_x f(x) \rightsquigarrow x^+ = x - \nabla^2 f(x)^{-1} \nabla f(x)$$

Assume f strongly convex, and both $\nabla f, \nabla^2 f$ are Lipschitz.

If $x^{(0)}$ near x^* then

$$x^{(k)} \rightarrow x^* \quad \text{and} \quad f(x^{(k)}) \rightarrow f^*$$

quadratically.

For global convergence use **damped Newton's method**:

$$x^+ = x - t \nabla^2 f(x)^{-1} \nabla f(x)$$

Newton's method for linearly-constrained optimization

For

$$\min_x f(x) \quad \text{subject to} \quad Ax = b$$

Newton's method

$$x^+ = x + tv$$

where

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} \nabla f(x) \\ Ax - b \end{bmatrix}.$$

The latter is precisely the root-finding Newton step for the KKT conditions of the above equality-constrained problem, namely

$$\begin{bmatrix} \nabla f(x) + A^T y \\ Ax - b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Barrier method

Consider the convex optimization problem

$$\begin{aligned} \min_x \quad & f(x) \\ \text{subject to} \quad & Ax = b \\ & h_i(x) \leq 0, \quad i = 1, \dots, m. \end{aligned}$$

Major challenge: figure out binding/non-binding constraints.
This complication occurs at the boundary of the feasible region.

Letting $C := \{x : h_i(x) \leq 0, \quad i = 1, \dots, m\}$, we can rewrite the above problem as

$$\begin{aligned} \min_x \quad & f(x) + I_C(x) \\ & Ax = b. \end{aligned}$$

Main idea of interior-point methods: approximate I_C with a **barrier** function for C to avoid the boundary of C and to make the problem amenable to Newton's method.

Log barrier function

Assume $h_1, \dots, h_m : \mathbb{R}^n \rightarrow \mathbb{R}$ are convex and twice differentiable.

The function

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

is called the **logarithmic barrier function** for the set

$$\{x : h_i(x) < 0, i = 1, \dots, m\},$$

which we assume is nonempty.

Approximate original problem with

$$\min_x \begin{array}{l} f(x) + \frac{1}{t}\phi(x) \\ Ax = b \end{array} \Leftrightarrow \min_x \begin{array}{l} tf(x) + \phi(x) \\ Ax = b \end{array}$$

where $t > 0$.

Outline

Today:

- Central path
- Properties and interpretations
- Barrier method
- Convergence analysis
- Feasibility methods

Log barrier calculus

For the log barrier function

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

we have

$$\nabla \phi(x) = - \sum_{i=1}^m \frac{1}{h_i(x)} \nabla h_i(x)$$

and

$$\nabla^2 \phi(x) = \sum_{i=1}^m \frac{1}{h_i(x)^2} \nabla h_i(x) \nabla h_i(x)^\top - \sum_{i=1}^m \frac{1}{h_i(x)} \nabla^2 h_i(x)$$

Central path

Given $t > 0$, consider the barrier problem defined above

$$\begin{aligned} \min_x \quad & tf(x) + \phi(x) \\ \text{subject to} \quad & Ax = b. \end{aligned}$$

Let $x^*(t)$ denote the solution to this above barrier problem. The **central path** is the set $\{x^*(t) : t > 0\}$.

Under suitable conditions, this set is a smooth path in \mathbb{R}^n and as $t \rightarrow \infty$, we have $x^*(t) \rightarrow x^*$, where x^* is a solution to our original problem.

Perturbed KKT conditions

KKT conditions for barrier problem:

$$t \nabla f(x^*(t)) - \sum_{i=1}^m \frac{1}{h_i(x^*(t))} \nabla h_i(x^*(t)) + A^T w = 0$$
$$Ax^*(t) = b, \quad h_i(x^*(t)) < 0, \quad i = 1, \dots, m$$

KKT conditions of the original problem:

$$\nabla f(x^*) + \sum_{i=1}^m u_i^* \nabla h_i(x^*) + A^T v^* = 0$$
$$Ax^* = b, \quad h_i(x^*) \leq 0, \quad u_i^* \geq 0, \quad i = 1, \dots, m$$
$$h_i(x^*) \cdot u_i^* = 0, \quad i = 1, \dots, m$$

Duality gap

By convexity we have

$$f(x^*(t)) - f(x^*) \leq \nabla f(x^*(t))^T (x^*(t) - x^*)$$

and

$$h_i(x^*(t)) - h_i(x^*) \leq \nabla h_i(x^*(t))^T (x^*(t) - x^*), \quad i = 1, \dots, m.$$

Thus the previous two sets of KKT conditions yield

$$f(x^*(t)) - f^* \leq \frac{m}{t}.$$

This is a useful stopping criterion.

Barrier method v.0

For $\epsilon > 0$ pick $t = m/\epsilon$ and solve

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

to get $f(x^*(t)) - f^* \leq \epsilon$.

This is not a good idea because the barrier problem is too difficult to solve.

The above approach aims to find a point **near the end** of the central path.

A better approach is to generate points **along** the central path.

Barrier method v.1

Solve a sequence of barrier problems

$$\begin{aligned} \min_x \quad & t f(x) + \phi(x) \\ \text{subject to} \quad & Ax = b \end{aligned}$$

for increasing values of t .

- Pick $t^{(0)} > 0$ and let $k := 0$
- Solve the barrier problem for $t = t^{(0)}$ to produce $x^{(0)} = x^*(t)$
- While $m/t > \epsilon$
 - ▶ Pick $t^{(k+1)} > t^{(k)}$
 - ▶ Solve the barrier problem at $t = t^{(k+1)}$, using Newton's method initialized at $x^{(k)}$, to produce $x^{(k+1)} = x^*(t)$

end while

Common update $t^{(k+1)} = \mu t^{(k)}$ for $\mu > 1$.

Centering step: the step that solves the barrier problem.

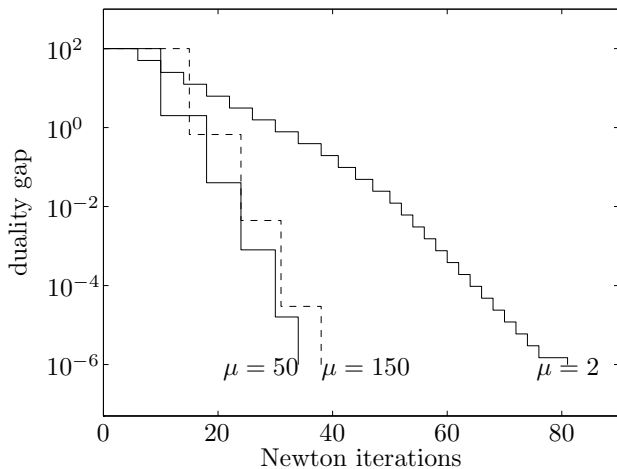
Considerations:

- **Choice of μ :** if μ is too small, then many outer iterations might be needed; if μ is too big, then Newton's method (each centering step) might take many iterations to converge.
- **Choice of $t^{(0)}$:** if $t^{(0)}$ is too small, then many outer iterations might be needed; if $t^{(0)}$ is too big, then the first Newton's solve (first centering step) might require many iterations to compute $x^{(0)}$.

Fortunately, the performance of the barrier method is often quite robust to the choice of μ and $t^{(0)}$ in practice.

The appropriate range for these parameters is scale dependent.

Example of a small LP in $n = 50$ dimensions, $m = 100$ inequality constraints (from B & V page 571):



Convergence analysis

Assume that we solve the centering steps exactly. The following result is immediate

Theorem: The barrier method after k centering steps satisfies

$$f(x^{(k)}) - f^* \leq \frac{m}{\mu^k t^{(0)}}$$

In other words, to reach a desired accuracy level of ϵ , we require

$$\frac{\log(m/(t^{(0)}\epsilon))}{\log \mu} + 1$$

centering steps with the barrier method (plus initial centering step).

Barrier method v.2

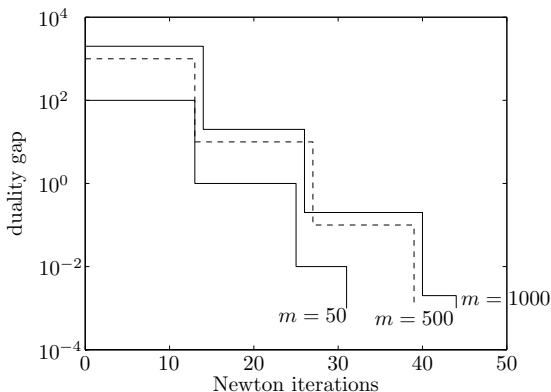
The previous algorithm generates points that are exactly on the central path. However, the central path is only a “means to an end”. There is no need to solve each problem exactly.

- Pick $t^{(0)} > 0$ and let $k := 0$.
- Solve the barrier problem for $t = t^{(0)}$ to produce $x^{(0)} \approx x^*(t)$
- While $m/t > \epsilon$
 - ▶ Pick $t^{(k+1)} > t^{(k)}$
 - ▶ Solve the barrier problem at $t = t^{(k+1)}$, using Newton's method initialized at $x^{(k)}$, to produce $x^{(k)} \approx x^*(t)$

Important issues (can be formalized):

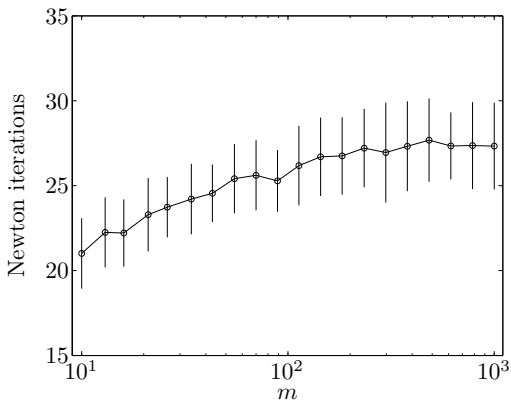
- How close should each approximation be?
- How many Newton steps suffice at each centering step?

Example of barrier method progress for an LP with m constraints (from B & V page 575):



Can see roughly linear convergence in each case, and logarithmic scaling with m

Seen differently, the number of Newton steps needed (to decrease initial duality gap by factor of 10^4) grows very slowly with m :



Note that the cost of a single Newton step does depend on the size of the problem.

Feasibility methods

We have implicitly assumed that we have a **strictly feasible** point for the first centering step, i.e., for computing $x^{(0)} = x^*$, solution of barrier problem at $t = t^{(0)}$.

This is a point x such that

$$h_i(x) < 0, \quad i = 1, \dots, m, \quad Ax = b$$

How to find such a feasible x ? By solving

$$\begin{aligned} \min_{x,s} \quad & s \\ \text{subject to} \quad & h_i(x) \leq s, \quad i = 1, \dots, m \\ & Ax = b. \end{aligned}$$

The goal is for s to be negative at the solution. This is known as a **feasibility method**. We can apply the barrier method to the above problem, since it is easy to find a strictly feasible starting point.

Note that we do not need to solve this problem to high accuracy. Once we find a feasible (x, s) with $s < 0$, we can **terminate early**.

An alternative is to solve the problem

$$\begin{aligned} \min_{x,s} \quad & \mathbf{1}^T s \\ \text{subject to} \quad & h_i(x) \leq s_i, \quad i = 1, \dots, m \\ & Ax = b, \quad s \geq 0. \end{aligned}$$

Previously s was the maximum infeasibility across all inequalities. Now each inequality has own infeasibility variable s_i , $i = 1, \dots, m$.

One advantage: when the original system is infeasible, the solution of the above problem will be informative. The **nonzero entries** of s will tell us which of the constraints cannot be satisfied.

References and further reading

- S. Boyd and L. Vandenberghe (2004), “Convex optimization”, Chapter 11
- A. Nemirovski (2004), “Interior-point polynomial time methods in convex programming”, Chapter 4
- J. Nocedal and S. Wright (2006), “Numerical optimization”, Chapters 14 and 19

A formal barrier method

A convex function $\phi : D \rightarrow \mathbb{R}$ defined on an open convex set $D \subseteq \mathbb{R}^n$ is a **self-concordant barrier with parameter ν** if

- ϕ is self-concordant
- For all $x \in D$ we have

$$\lambda(x)^2 = \nabla\phi(x) (\nabla^2\phi(x))^{-1} \nabla\phi(x) \leq \nu.$$

Consider the problem

$$\begin{aligned} \min_x \quad & c^\top x \\ \text{subject to} \quad & x \in \bar{D} \end{aligned}$$

Approximate with

$$\min_x \quad tc^\top x + \phi(x).$$

For convenience let $\phi_t(x) := tc^\top x + \phi(x)$ and let $\lambda_t(x)$ denote the corresponding Newton decrement.

Key observation: for $t^+ > t$

$$\lambda_{t^+}(x) \leq \frac{t^+}{t} \lambda_t(x) + \left(\frac{t^+}{t} - 1 \right) \sqrt{\nu}.$$

Theorem

If $\lambda_t(x) \leq \frac{1}{9}$ and $\frac{t^+}{t} \leq 1 + \frac{1}{8\sqrt{\nu}}$ then $\lambda_{t^+}(x^+) \leq \frac{1}{9}$ for

$$x^+ = x - (\nabla^2 \phi_{t^+}(x))^{-1} \nabla \phi_{t^+}(x).$$

Consequently, if we start with $x^{(0)}, t^{(0)}$ such that $\lambda_{t^{(0)}}(x^{(0)}) < \frac{1}{9}$ and choose $\mu := 1 + \frac{1}{8\sqrt{\nu}}$ in the barrier method, one Newton iteration suffices at each centering step.