Log Barrier Method

The inequality-constrained optimization problem is stated as

$$\min_x f(x)$$
$$\text{subject to } c(x) \geq 0,$$

where both $f(x)$ and $c(x)$ are assumed to be continuous.

Log barrier method is to find an unconstrained minimizer of a composite function that reflects the original objective function as well as the presence of constraints. The logarithmic barrier function is defined as

$$B(x, \mu) = f(x) - \mu \sum_{i=1}^{m} \ln c_i(x),$$

where μ is a positive scalar, the barrier parameter. $B(x, \mu)$ retains the smoothness properties of $f(x)$ and $c(x)$ as long as $c(x) > 0$. For very small $\mu > 0$, $B(x, \mu)$ acts like $f(x)$ except close to points where any constraint is zero. Intuition suggests that minimizing $B(x, \mu)$ for a sequence of positive μ values converging to zero will cause the unconstrained minimizers of $B(x, \mu)$ to converge to a local constrained minimizer of the original problem. The gradient of the barrier function, denoted by $\nabla B(x, \mu)$, is

$$\nabla B(x, \mu) = \nabla f(x) - \sum_{i=1}^{m} \frac{\mu}{c_i(x)} \nabla c_i(x).$$

An unconstrained minimizer will be denoted by x_μ, and it will be proven later that $c(x_\mu) > 0$. By the optimality conditions for unconstrained optimization (P591 Lemma A7), it must hold that $\nabla B(x_\mu, \mu) = 0$ when $\nabla B(x, \mu)$ is twice-continuously differentiable. This leads to that

$$\nabla f(x_\mu) = \sum_{i=1}^{m} \frac{\mu}{c_i(x)} \nabla c_i(x).$$

This implies that the objective gradient at x_μ is a positive linear combination of the constraint gradients. The coefficients in the linear combination are called the barrier multiplier (analogy with Lagrange multipliers), denoted by λ_μ. Formally, λ_μ is defined as

$$\lambda_\mu \triangleq \mu \cdot / c(x),$$

with $\lambda_\mu > 0$. Thus, we have

$$c(x_\mu) \cdot \lambda_\mu = \mu.$$

This relationship betwwen the barrier multipliers, constraint values, and the barrier parameter, called perturbed complementarity, is analogous as $\mu \to 0$ to the complementarity condition $c(x) \cdot \lambda = 0$ that holds at a KKT point.
??
Consider this inequality constrained optimization problem. Here is a theorem, called existence of compact enclosing set theorem, needed for the local convergence theorem. $\mathcal{N}$ denote the set of all local constrained minimizers with objective function value f^*, and assume that f^* has been chosen so that $\mathcal{N}$ is non-empty. Assume further that the set $\mathcal{N}^* \subseteq \mathcal{N}$ is a nonempty compact isolated subset of $\mathcal{N}$. Then there exist a compact set S such that $\mathcal{N}^*$ lies in $int(S) \cap \mathcal{F}$ and $f(y) > f^*$ for any feasible point y in S but not in $\mathcal{N}^*$. Every point x^* in $\mathcal{N}^*$ thus has the property that $f(x^*) = f^* = minf(x)$ for all $x \in S \cap \mathcal{F}$, where $\mathcal{F}$ is the feasible region.
??

THEOREM: Local convergence for barrier methods
Consider the problem of minimizing $f(x)$ subject to $c(x) \geq 0$, where $f(x)$ and $c(x)$ are continuous. Let $\mathcal{F}$ denote the feasible region, let $\mathcal{N}$ denote the set of minimizers with objective function value $f^* = \min f(x)$, and assume that $\mathcal{N}$ is non-empty. Let $\{\mu_\kappa\}$ be a strictly decreasing sequence of positive barrier parameters such that $\lim_{k\to\infty} \mu_k = 0$. Assume that

(a) there exists a non-empty compact set $\mathcal{N}^*$ of local minimizers that is an isolated subset of $\mathcal{N}$;

(b) at least one point in $\mathcal{N}^*$ is in the closure of strict$(\mathcal{F})$, i.e., there is at lease one point in $\mathcal{N}^*$ that is strictly feasible or the limiting point of $\mathcal{F}$.

Then the following results hold:

(i) there exists a compact set S such that $\mathcal{N}^* \subset \text{int}(S)$ and such that , for any feasible point $\bar{x}$ in S but not in $\mathcal{N}^*$, $f(\bar{x}) > f^*$;

(ii) for all sufficiently small μ_k, there is an unconstrained minimizer y_k of the barrier function $B(x, \mu_k)$ in strict$(\mathcal{F}) \cap \text{int}(S)$, with

$$B(y_k, \mu_k) = \min\{B(x, \mu_\kappa) : x \in strict(\mathcal{F}) \cap S)\}.$$

Thus $B(y_\kappa, \mu_\kappa)$ is the smallest value of $B(x, \mu_\kappa)$ for any $x \in \text{strict}(\mathcal{F}) \cap S$.

(iii) any sequence of these unconstrained minimizers $\{y_k\}$ of $B(x, \mu_\kappa)$ has at least one convergent subsequence;

(iv) the limit point x_∞ of any convergent subsequence $\{x_k\}$ of the unconstrained minimizers $\{y_k\}$ defined in (ii) lies in $\mathcal{N}^*$.

(v) for the convergent subsequences $\{x_k\}$ of part (iv)

$$\lim_{k\to\infty} f(x_\kappa) = f^* = \lim_{k\to\infty} B(x_\kappa, \mu_\kappa).$$

THEOREM: Properties of the central path/barrier trajectory

Consider the problem of minimizing $f(x)$ subject to $c(x) \geq 0$. Let $\mathcal{F}$ denote the feasible region, and assume that the set strict($\mathcal{F}$) of strictly feasible points is non-empty. Let x^* be a local constrained minimizer, with g^* denoting $g(x^*)$, J^* denoting $J(x^*)$, and so on, and let $\mathcal{A}$ denote $\mathcal{A}(x^*)$. Assume that the following sufficient optimality conditions hold at x^*:

(a) x^* is a KKT point, i.e., there exists a nonempty set $\mathcal{M}_\lambda$ of Lagrange multipliers λ satisfying

$$\mathcal{M}_\lambda = \{\lambda : g^* = J^{*T}\lambda, \lambda \geq 0, \text{and } c(x^*)\cdot\lambda = 0\}$$

(b) the MFCQ (a condition on the constraints) holds at x^*, i.e., there exists p such that $J^*_{\mathcal{A}}p > 0$, where $J^*_{\mathcal{A}}$ denotes the Jacobian of the active constraints at x^*; and

(c) there exists $\omega > 0$, such that $p^T H(x^*, \lambda)p \geq w\|p\|^2$ for all $\lambda \in \mathcal{M}_\lambda$ and all nonzero p satisfying $g^{*T}p = 0$ and $J^*_{\mathcal{A}}p \geq 0$, where $H(x^*, \lambda)$ is the Hessian of the Lagrangian (2.11). $H(x, \lambda) \triangleq \nabla^2_{xx}L(x, \lambda) = \nabla^2 f(x) - \sum_{i=1}^m \lambda_i \nabla^2 c_i(x)$

Assume that a logrithmic barrier method is applied in which μ_k converges monotonically to zero as $k \to \infty$. Then,

(i) there is at least one subsequence of unconstrained minimizers of the barrier function $B(x, \mu_k)$ converging to x^*;

(ii) let $\{x^k\}$ denote such a convergent subsequence, with the obvious notation that c_i^k denotes $c_i(x^k)$, and so on. Then the sequence of barrier multipliers $\{\lambda^k\}$, whose i-th component is μ_k/c_i^k, is bounded;

(iii) $\lim_{k\to\infty}\lambda^k = \bar{\lambda} \in \mathcal{M}_\lambda$

If, in addition, strict complementarity holds at x^*, i.e, there is a vector $\lambda \in \mathcal{M}_\lambda$ such that $\lambda_i > 0$ for all $i \in \mathcal{A}$, then

(iv) $\bar{\lambda}_{\mathcal{A}} > 0$;

(v) for sufficiently large k, the Hessian matrix $\nabla^2 B(x^k, \mu_k)$ is positive defnite;

(vi) a unique, continuously differentiable vector function $x(\mu)$ of unconstrained minimizers of $B(x, \mu)$ exisits for positive μ in a neighborhood of $\mu = 0$; and

(vii) $\lim_{\mu\to 0_+} x(\mu) = x^*$

Problem

$$\min_{\tau} \iint -\operatorname{sgn}(v)\, u\, f_Y(u, v; \tau)\; du\, dv$$

$$\text{subject to } \kappa - \iint \operatorname{sgn}(v)\, w\, f_Z(w, v; \tau)\; dw\, dv \geq 0$$

Log-barrier formation

$$B(\tau,\mu)=\iint -\text{sgn}(v)uf_Y(u,v;\tau)\,du\,dv-\mu\text{ln}\left[\kappa-\iint \text{sgn}(v)wf_Z(w,v;\tau)\,dw\,dv\right],$$

and

$$\nabla B(\tau,\mu)=\iint -\text{sgn}(v)u\nabla f_Y(u,v;\tau)\,du\,dv+\mu\frac{\iint \text{sgn}(v)w\nabla f_Z(w,v;\tau)\,dw\,dv}{\kappa-\iint \text{sgn}(v)wf_Z(w,v;\tau)\,dw\,dv},$$

where μ is a sequence of decreasing positive constants converging to zero. Note that ∇ is the first order derivative with respect to τ.

We denote an unconstrained minimizer of $B(\tau,\mu)$ as $\tau(\mu)$ or τ_μ. $\tau(\mu)$ is used when we need to emphasize it as a function of the barrier parameter μ, and $\tau(\mu)$ is used for short notation embedded in equations. It can be proven that the constraint is strictly satisfied, i.e., $\kappa-\iint \text{sgn}\,(v)\,w\,f_Z\,(w,v;\tau_\mu)\;dw\,dv\;>\;0$. Assume that $\nabla B(\tau,\mu)$ is twice-continuously differentiable, it must hold that $\nabla B(\tau_\mu,\mu)=0$, which means that

$$\iint \text{sgn}\,(v)\,u\,\nabla f_Y\,(u,v;\tau_\mu)\;du\,dv=\mu\frac{\iint \text{sgn}(v)w\nabla f_Z(w,v;\tau)\,dw\,dv}{\kappa-\iint \text{sgn}(v)wf_Z(w,v;\tau_\mu)\,dw\,dv}$$

The barrier multiplier, the coefficient in that linear relationship, is denoted by λ_μ and is defined as

$$\lambda_\mu\triangleq\frac{\mu}{\kappa-\iint \text{sgn}(v)wf_Z(w,v;\tau_\mu)\,dw\,dv}$$

This relationship can be written as

$$\lambda_\mu\left[\kappa-\iint \text{sgn}(v)wf_Z(w,v;\tau_\mu)\,dw\,dv\right]=\mu.$$

This relationship between the barrier multiplier, constraint value, and the barrier parameter, called perturbed complementarity, is analogous as $\mu\to 0$ to the complementarity condition $c(x^*)\lambda^*=0$ that holds at a KKT point.

Optimality conditions for the central path/barrier trajectory

We display the optimality conditions for the central path/barrier trajectory. Consider the problem stated above. Let $\mathcal{F}$ denote the feasible region, and assume that the set strict$(\mathcal{F})$ of strictly feasible points is non-empty.Let τ^* be a local constrained minimizer, with g^* denoting $g(\tau^*)=\nabla f(\tau^*)$, J^* denoting $J(\tau^*)=\nabla c(\tau^*)$, and so on, and let $\mathcal{A}$ denote $\mathcal{A}(\tau^*)$. Assume theat the following sufficient optimality conditions hold at τ^*:

(a) τ^* is a KKT point, i.e., there exists a nonempty set $\mathcal{M}_\lambda$ of Lagrange multipliers λ satisfying

$$\mathcal{M}_\lambda = \{\lambda : g^* = J^{*T}\lambda, \lambda \geq 0, \text{and } c(\tau^*) \cdot \lambda = 0\}$$

(b) the MFCQ (a condition on the constraints) holds at τ^*, i.e., there exists p such that $J^*_{\mathcal{A}}p > 0$, where $J^*_{\mathcal{A}}$ denotes the Jacobian of the active constraints at τ^*; and

(c) there exists $\omega > 0$, such that $p^T H(\tau^*, \lambda)p \geq w\|p\|^2$ for all $\lambda \in \mathcal{M}_\lambda$ and all nonzero p satisfying $g^{*T}p = 0$ and $J^*_{\mathcal{A}}p \geq 0$, where $H(x^*, \lambda)$ is the Hessian of the Lagrangian (2.11). $H(\tau, \lambda) \triangleq \nabla^2_{\tau\tau} L(\tau, \lambda) = \nabla^2 f(\tau) - \sum_{i=1}^m \lambda_i \nabla^2 c_i(\tau)$

Assume that a logrithmic barrier method is applied in which μ_k converges monotonically to zero as $k \to \infty$. Then,

(i) there is at least one subsequence of unconstrained minimizers of the barrier function $B(\tau, \mu_k)$ converging to τ^*;

(ii) let $\{\tau^k\}$ denote such a convergent subsequence, with the obvious notation that c_i^k denotes $c_i(\tau^k)$, and so on. Then the sequence of barrier multipliers $\{\lambda^k\}$, whose i-th component is μ_k / c_i^k, is bounded;

(iii) $\lim_{k\to\infty} \lambda^k = \bar{\lambda} \in \mathcal{M}_\lambda$

If, in addition, strict complementarity holds at τ^*, i.e, there is a vector $\lambda \in \mathcal{M}_\lambda$ such that $\lambda_i > 0$ for all $i \in \mathcal{A}$, then

(iv) $\bar{\lambda}_{\mathcal{A}} > 0$;

(v) for sufficiently large k, the Hessian matrix $\nabla^2 B(\tau^k, \mu_k)$ is positive defnite;

(vi) a unique, continuously differentiable vector function $\tau(\mu)$ of unconstrained minimizers of $B(\tau, \mu)$ exisits for positive μ in a neighborhood of $\mu = 0$; and

(vii) $\lim_{\mu\to 0_+} \tau(\mu) = \tau^*$