Chap 17: Interior point methods.

Let $P = \{\vec{x} : A\vec{x} \le \vec{b}, \vec{x} \ge \vec{0}\}$ be feasible region of an LP problem.

- Simplex moves along the "boundary" of P, to the solution where nonbasic variable are 0.

- Interior-point methods make a path to the solution in the interior of P

(P)

$$\max \vec{c}^T\vec{x}$$
$$\text{s.t. } A\vec{x} + \vec{w} = \vec{b}$$
$$\vec{x}, \vec{w} \ge \vec{0}$$

(D)

$$\min \vec{b}^T\vec{y}$$
$$\text{s.t. } A^T\vec{y} - \vec{z} = \vec{c}$$
$$\vec{y}, \vec{z} \ge \vec{0}$$

Idea: Restate (P) and (D) to eliminate the constraint $\vec{x}, \vec{y}, \vec{w}, \vec{z} \ge 0$, but still avoid values ≤ 0 for these. Will achieve this by using a logarithmic function

Barrier problem

$$(P_\mu): \quad \max \vec{c}^T\vec{x} + \mu \sum_{j=1}^{n} \log x_j + \mu \sum_{i=1}^{m} \log w_i$$
$$\text{s.t. } A\vec{x} + \vec{w} = \vec{b}$$

Remarks:

- (P_μ) is an approximation to (P), the constraint $\vec{x}, \vec{w} \ge \vec{0}$ are implicit in the objective.
- As $\mu \to 0^+$ the objective of (P_μ) becomes more and more like that of (P).
- (P_μ) is nonlinear.

Goal: Show that (P_μ) has a unique solution $\vec{x}(\mu)$, so that $\vec{x}(\mu) \to \vec{x}^*$ as $\mu \to 0^+$, where $\vec{x}^*$ is the optimal solution to (P).

Theorem 5.10.5 in FVLA (MAT1110)

Suppose $U \subseteq \mathbb{R}^n$ is open, and that f, g_i ($i = 1, \dots, m$) have cont. part. der. in U.
Let $b_1, \dots, b_m \in \mathbb{R}$, and assume that $\vec{x}^*$ is a local max/min for f on the set

$$S = \{ \vec{x} \in \mathbb{R}^n : g_i(\vec{x}) = b_i \ , i = 1, \dots, m \},$$

and that $\nabla g_1(\vec{x}^*), \dots, \nabla g_m(\vec{x}^*)$ are linearly independent.
Then there exist constants $\lambda_1, \dots, \lambda_m$ so that

$$(*) \quad \nabla f(x^*) = \lambda_1 \nabla g_1(x^*) + \dots + \lambda_m \nabla g_m(x^*)$$

(necessary condition for optimality)

The $\lambda_1, \dots, \lambda_m$ are called Lagrange multipliers.
Any x^* satisfying $(*)$ is called a critical point.

So: For a local max/min x^*, either 1. The $\nabla g_i(x^*)$ are lin. dep.
2. x^* is a critical point

Consider the Lagrange function: $L(\vec{x}, \vec{\lambda}) = f(\vec{x}) - \sum_{i=1}^{m} \lambda_i g_i(\vec{x})$

with g_i replaced with $g_i(\vec{x}) - b$ (so that constraints take the form $g_i(\vec{x}) = 0$.

Note that $(*)$ is the same as $\nabla_x L(x^*, \vec{\lambda}) = \vec{0}$

$g_i(\vec{x}) = \vec{0}$ is the same as $\nabla_\lambda L(x^*, \vec{\lambda}) = \vec{0}$

So, $\nabla L(x^*, \lambda) = \vec{0}$ captures both the constraints $g_i(\vec{x}) = 0$, and $(*)$

Also called first order optimality conditions.

Sufficient conditions for optimality:

Define $H_f(\vec{x}) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right]_{i,j}$ (symmetric $n \times n$ matrix).

Theorem 17.1 Second order optimality conditions

If the g_i are linear functions, and x^* is a critical point, then x^* is a local maximum in the constrained problem if

$$\vec{z}^T Hf(x^*)\vec{z} < 0 \text{ for all } \vec{z} \neq \vec{0} \text{ satisfying}$$
$$\vec{z}^T \nabla g_i(x^*) = 0 \ \forall i$$

A $\vec{z}$ satisfying $\vec{z}^T \nabla g_i(x^*)$ $\forall i$ is called a feasible direction

$$g_i(x) = a + \nabla g_i(x^*)x, \quad g_i(x+z) = \underbrace{a + \nabla g_i(x^*)(x+z)}_{a + \nabla g_i(x^*)x + \nabla g(x^*)z}$$

must be 0 for feasible dir. $\vec{z}$.

Proof: Second order Taylor says:

$$f(x^*+z) = f(x^*) + \underbrace{\nabla f(x^*)z}_{0} + \underbrace{\frac{1}{2} z^T Hf(x^*) z}_{<0} + \underbrace{O(\|z\|^2)}_{k(z)}$$

means $\frac{k(z)}{\|z\|^2} \to 0$ as $z \to 0$

To preserve feasibity, $\vec{z}$ must be a feasible direction.

Since $\vec{x}$ is a critical and $\vec{z}$ is a feasible direction:

$$\nabla f(x^*)\vec{z} = (\lambda_1 \nabla g_1(x^*) + \cdots \lambda_m \nabla g_m(x^*))\vec{z}$$
$$= \underbrace{\lambda_1 \nabla g_1(x^*)\vec{z}}_{0} + \cdots + \underbrace{\lambda_m \nabla g_m(x^*)\vec{z}}_{0}$$
$$= 0$$

This, and that $z^T Hf(x^*) z < 0$ for all feasible directions, and that $|\frac{1}{2} z^T Hf(x^*) z|$ dominates $O(\|z\|^2)$, imply that $f(x^* + z) < f(x^*)$, so that x^* is local max ∎

Note: If $z^T Hf(x^*) z < 0$ for all x, then x^* is also a unique global maximum! Following version of Taylor's theorem is needed for this:

$$f(x^* + h) = f(x^*) + \underbrace{\nabla f(x^*)^T h}_{0} + \underbrace{\frac{1}{2} h^T f(x + th) h}_{0 \le t \le 1} < f(x^*)$$

$$< 0$$

Lagrange method applied to barrier problem

$$(P_\mu): \quad \max \; \underbrace{c^T x + \mu \sum_{j=1}^{n} \log(x_j) + \mu \sum_{i=1}^{m} \log(w_i)}_{f(x,w)}$$

$$\text{s.t.} \quad Ax + w = b \qquad g_i(x,w) = (Ax + w - b)_i$$

Lagrange function is

$$L(x, w, y) = f(x,w) - \sum_{i=1}^{m} y_i \, g_i(x,w)$$

$$= \underbrace{c^T x + \mu \sum_{j=1}^{n} \log x_j + \mu \sum_{i=1}^{m} \log w_i}_{f(x,w)} - y^T (Ax + w - b)$$

First order optimality conditions

$$\frac{\partial L}{\partial x_j} = c_j + \frac{\mu}{x_j} - \sum_{i=1}^{m} y_i a_{ij} = 0$$

$$\frac{\partial L}{\partial w_i} = \frac{\mu}{w_i} - y_i = 0$$

$$\frac{\partial L}{\partial y_i} = (Ax + w - b)_i = 0$$

Write X, W for the diagonal matrices with $\vec{x}$, $\vec{w}$ on the diagonal, and $\vec{e}$ the vector of ones.

$$\begin{aligned} c + \mu X^{-1}\vec{e} - A^T\vec{y} &= 0 \\ \mu W^{-1}\vec{e} - \vec{y} &= 0 \\ b - A\vec{x} - \vec{w} &= 0 \end{aligned}$$

$$\vec{z} = \mu X^{-1}\vec{e}$$

$$\begin{aligned} A^T\vec{y} - \vec{z} &= c \\ \vec{y} &= \mu W^{-1}\vec{e} \\ A\vec{x} + \vec{w} &= b \end{aligned}$$

$$\begin{aligned} Ax + w &= b \\ A^T y - z &= c \\ X\cdot\vec{z} &= \mu\vec{e} \\ W\cdot\vec{y} &= \mu\vec{e} \end{aligned}$$

primal feasibility

dual feasibility

μ-complementarity.

Second order optimality conditions (uniqueness of solution to barrier problem.

With $f(x,w) = c^T x + \mu \sum_{j=1}^{n} \log x_j + \mu \sum_{i=1}^{m} \log w_i$, we have

First order

$$\begin{cases} \dfrac{\partial f}{\partial x_j} = c_i + \dfrac{\mu}{x_j} \\ \dfrac{\partial f}{\partial w_i} = \dfrac{\mu}{w_i} \end{cases}$$

Second order:

$$\frac{\partial^2 f}{\partial x_j^2} = -\frac{\mu}{x_j^2}$$

$$\frac{\partial^2 f}{\partial w_i^2} = -\frac{\mu}{w_i^2}$$

all other partial derivatives are 0.

Thus, Hf is diagonal with negative diagonal entries,

$$\Downarrow$$

$z^T Hf\, z < 0$ always.

From the note following the proof of theorem 17.1:
A critical point is also a unique global maximum.

Existence of solution to barrier problem

Theorem 17.2: The barrier problem has a solution

$\Updownarrow$

Both primal and dual feasible regions have nonempty interiors.

Proof: $\Downarrow$: is trivial

$\Uparrow$: Assume $\begin{matrix}(\bar{x},\bar{w})\\(\bar{y},\bar{z})\end{matrix}$ interior to $\begin{matrix}\text{primal}\\\text{dual}\end{matrix}$ feasible region

If (x,w) is primal feasible:

$$\bar{z}^T x + \bar{y}^T w = (A^T\bar{y} - c)^T x + \bar{y}^T(b - Ax)$$
$$= b^T\bar{y} - c^T x$$

Thus: $c^T x = -\bar{z}^T x - \bar{y}^T w + b^T\bar{y}$

Barrier function: $f(x,w) = c^T x + \mu\sum_{j=1}^{n}\log x_j + \mu\sum_{i=1}^{m}\log w_i$

$$= \sum_{j=1}^{n}(-\bar{z}_j x_j + \mu\log x_j) + \sum_{i=1}^{m}(-\bar{y}_i w_i + \mu(\log w_i)) + b^T\bar{y}$$

Both are on the form $h(v) = -av + \mu\log v$, $a>0$, $0<v<\infty$

$h'(v) = -a + \frac{\mu}{v}$, $=0$ if $v = \frac{\mu}{a}$

$h''(v) = -\frac{\mu}{v^2} < 0$,

so that $v = \frac{\mu}{a}$ is max for h

It follows: $\{(x,w): f(x,w) \geq \delta\}$ is bounded for any δ

Set $\delta = \bar{f} = f(\bar{x},\bar{w})$, define $\bar{P} = \{(x,w): Ax + w = b,\ x,w \geq 0,\ f(x,w) \geq \bar{f}\}$

We have that $\bar{P} = \{(x,w): Ax + w = b\}$ (closed)

$\cap\{(x,w): f(x,w) \geq \bar{f}\}$ (closed, bounded)

$\cap\{(x,w): x,w \geq 0\}$ (closed)

Thus, $\bar{P}$ is closed and bounded. $\bar{P}$ is also nonempty ($(\bar{x},\bar{w}) \in \bar{P}$)

f thus has a supremum on P, and therefore also on

$\{(x,w): Ax+w=b,\ x,w \geq 0\}$

$\bar{P}$ is this set, with small f-values cut out ■

Exercise 10.7: The dual feasible region has an interior point when the primal feasible region is bounded.

Corollary 17.3: If the primal feasible region is bounded and has interior points, then for each μ there exists a unique solution $(x(\mu), w(\mu))$ to the barrier problem (similarly for $(y(\mu), z(\mu))$)

The path $(x(\mu), w(\mu), y(\mu), z(\mu))$ is also called the primal/dual central path

primal-dual path-following method: One computes a sequence

$$\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(k)} \to 0$$

For each k one approximately solves the barrier problem with parameter $\mu^{(k)}$ using Newton's method.

$p(\mu^{(k)}) \to$ optimal primal/dual solution.