Chapter 17

Let P be the feasible solutions of an LP problem

- Simplex algorithm moves along the boundary of P (nonbasic variables are zero)
- Interior point methods make a path to the solution in the interior of P.

  Often faster algorithms

(P) $\max \vec{c}^T\vec{x}$ s.t. $A\vec{x} \le \vec{b}$, $\vec{x} \geq \vec{0}$	(D) $\min \vec{b}^T\vec{y}$ s.t. $A^T\vec{y} \geq \vec{c}$, $\vec{y} \geq \vec{0}$
with slack variables: $\max \vec{c}^T\vec{x}$ s.t. $A\vec{x} + \vec{w} = \vec{b}$, $\vec{x}, \vec{w} \geq \vec{0}$	$\min \vec{b}^T\vec{y}$ s.t. $A^T\vec{y} - \vec{z} = \vec{c}$, $\vec{y}, \vec{z} \geq 0$

Idea: Rewrite problems to eliminate $\vec{x}, \vec{w} \geq \vec{0}$, $\vec{y}, \vec{z} \geq 0$, i.e., make variables unconstrained, but still avoid values ≤ 0.

Will achieve this using a logarithmic barrier function.

Barrier problem:

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

- (P_μ) is an approximation to (P), with variable constraints observed in the objective
- As $\mu \to 0$, the objective function for (P_μ) becomes more and more like that of (P) (fig 17.1 in the book).
- (P_μ) is nonlinear.

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

Theorem 5.10.5 in FVLA (MAT 1110)

Suppose $U \subseteq \mathbb{R}^n$ is open, $f, g_i, U \to \mathbb{R}$ have cont. part. der.
Let $b_1, \dots, b_m \in \mathbb{R}$
Assume that x^* is a local max/min for f on the set

$$S = \{ x \in \mathbb{R}^n \mid g_i(x) = b_i \text{ for all } i \},$$

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

$$(*) \quad \nabla f(x^*) = \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*).$$

Necessary optimality condition.

Alternatively expressed in terms of the Lagrange function as follows:

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

(replaced $g_i \to g_i - b_i$, so that constraints are $g_i(x) = 0$)

1) $(*)$ is the same as $\nabla_x L(x^*, \lambda) = \vec{0}$
x^* is then called a critical point for the constrained problem.

2) $g_i(x^*) = 0$ is the same as $\underbrace{\nabla_y L(x^*, \lambda)}_{(g_1(x), \dots, g_m(x))} = \vec{0}$

$\nabla L(x^*, \lambda) = \vec{0}$ are called first order optimality conditions.

So, for a max/min it is necessary that either
1) The $\nabla g_i(x^*)$ linearly dependent
2) x^* is a critical point

Sufficient optimality conditions

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

Theorem 17.1

If the g_i are linear functions, and x^* is a critical point, then x^* is a local maximum in the constrained problem if $\vec{z}^T H_f(x^*) \vec{z} < 0$ for all $\vec{z} \neq \vec{0}$ which satisfy $\underbrace{\vec{z}^T \nabla g_i(x^*) = 0 \text{ for all } i}_{\vec{z} \text{ feasible direction}}$

Second order optimality condition

Proof: Second order Taylor formula says

$$f(\vec{x}^* + \vec{z}) = f(\vec{x}^*) + \nabla f(\vec{x}^*)^T \vec{z} + \frac{1}{2} \vec{z}^T H_f(\vec{x}^*) \vec{z} + \overbrace{O(\|z\|^2)}^{\text{means that } \lim_{z \to 0} \frac{K(z)}{\|z\|^2} = 0,\ K(z)}$$

To preserve feasibility, $\vec{z}$ must be chosen so that $x^* + z$ also satisfies the constraints.

No change in g_i $\Leftrightarrow$ change direction $\vec{z}$ perpendicular to ∇g_i
$\Leftrightarrow z^T \nabla g_i(x^*) = 0$ for all i

Since x^* is a critical point:

$$\nabla f(x^*)^T \vec{z} = \left(\sum_{i=1}^m \lambda_i \nabla g_i(x^*) \right)^T \vec{z} = \sum_{i=1}^m \lambda_i \nabla g_i(x^*)^T \vec{z} = 0$$

This, and that $\vec{z}^T H_f(x^*) \vec{z} < 0$ for all feasible directions $\vec{z}$, imply that

$$f(\vec{x}^* + \vec{z}) < f(x^*) \qquad \left(\underbrace{\tfrac{1}{2} \vec{z}^T H_f(x^*) \vec{z}}_{\text{min eigenvalue of } H_f(\vec{x})} \text{ dominates } O(\|z\|^2) \right)$$

So x^* is a local maximum.

Note: if $\vec{z}^T H_f(\vec{x}) \vec{z} < 0$ for all $\vec{x}$, then $\vec{x}^*$ is also a global maximum: Needs following for this:

$$f(\vec{x} + \vec{h}) = f(\vec{x}) + \nabla f(\vec{x})^T \vec{h} + \underbrace{\frac{1}{2} \vec{h}^T f(\vec{x} + t\vec{h}) \vec{h}}_{<0 \text{ always}}$$

$t \in (0,1)$

Lagranges method applied to the barrier problem

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

$$\text{s.t} \quad \underbrace{A\vec{x} + \vec{w} = \vec{b}}$$

$$g_i(\vec{x},\vec{w}) = (\vec{b} - A\vec{x} - \vec{w})_i = 0$$

$$G(\vec{x},\vec{w}) = \vec{b} - A\vec{x} - \vec{w} = \vec{0}$$

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

$$= \vec{c}^T\vec{x} + \mu\sum_{j=1}^{n}\log x_j + \mu\sum_{i=1}^{m}\log w_i - \vec{y}^T(\vec{b} - A\vec{x} - \vec{w})$$

First order optimality conditions:

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

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

$$\frac{\partial L}{\partial y_i} = b_i - \sum_{j=1}^{n} a_{ij} x_j - w_i = 0$$

write X for the diagonal matrix with the vector $\vec{x}$ on the diagonal
$\vec{e}$ is vector with all ones

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

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

rewrite to

$$A\vec{x} + \vec{w} = \vec{b}$$
$$A^T\vec{y} - \vec{z} = \vec{c}$$
$$X\vec{z} = \mu\vec{e}$$
$$W\vec{y} = \mu\vec{e}$$

interpretation

$\vec{x}$ primal feasible

$\vec{y}$ dual feasible

$\left.\begin{aligned} x_i z_i &= \mu \\ w_i y_i &= \mu \end{aligned}\right\}$ μ-complementarity. only difference from LP.

Second order optimality conditions (uniqueness of solution to barrier p.)

Recall that $f(\vec{x},\vec{w}) = \vec{c}^{\,T}\vec{x} + \mu \sum_j \log x_j + \mu \sum_i \log w_i$

First order:
$$\left\{\begin{array}{l} \dfrac{\partial f}{\partial x_j} = c_j + \dfrac{\mu}{x_j} \\[2ex] \dfrac{\partial f}{\partial w_i} = \dfrac{\mu}{w_i} \end{array}\right.$$

Second order:
$$\left\{\begin{array}{l} \dfrac{\partial^2 f}{\partial x_j^2} = -\dfrac{\mu}{x_j^2} \\[2ex] \dfrac{\partial^2 f}{\partial w_i^2} = -\dfrac{\mu}{w_i^2} \\[2ex] \text{others are } 0 \end{array}\right\}$$

$H^2 f$ is diagonal, with negative elements on the diagonal, so

$\vec{z}^{\,T} H^2 f \,\vec{z} < 0$ for all $(\vec{x},\vec{w})$.

From theorem 17.1 it follows that barrier function can have at most one critical point, and that this is a global maximum if it exists.

Existence of solution to the barrier problem

Theorem 17.2 Barrier problem has a solution if and only if both the primal feasible and the dual feasible region have a nonempty interior

Proof: We only show $\Leftarrow$

Assume $(\bar{x},\bar{w})$ is interior to primal feasible region

—"— $(\bar{y},\bar{z})$ —"— dual —"—

We have that $A\bar{x}+\bar{w} = b \qquad \bar{x},\bar{w},\bar{y},\bar{z} > 0$

$A^T\bar{y} - \bar{z} = c$

Assume that (x,w) is primal feasible:

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

It follows that $c^T\bar{x} = -\bar{z}^T x - \bar{y}^T w + b^T \bar{y}$

Barrier function: $f(x,w) = c^T x + \mu \sum_j \log x_j + \mu \sum_i \log w_i$

$$= \sum_j \left(-\bar{z}_j x_j + \mu \log x_j\right) + \sum_i \left(-\bar{y}_i w_i + \mu \log w_i\right) + b^T\bar{y}$$

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

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

$h''(v) = -\frac{\mu}{v^2} < 0$, so that $v = \frac{\mu}{a}$ is a maximum

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

Set $\delta = \bar{f} = f(\bar{x},\bar{w})$, and define

$$\bar{P} = \{(x,w) : Ax + w = b, x \geq 0, w \geq 0\} \quad \text{closed}$$
$$\cap \{(x,w) \quad x>0, w>0, f(x,w) \geq \bar{f}\} \quad \text{closed}$$

$\bar{P}$ is closed and bounded (i.e., $\bar{P}$ is compact)

$\bar{P}$ is also nonempty (contains $(\bar{x},\bar{w})$)

f thus has a supremum on $\bar{P}$, and therefore also on

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

as desired ■

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

It follows that:

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

The path $p(\mu) = \{(x(\mu), w(\mu), y(\mu), z(\mu))\}$ is 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 using Newtons method. $p(\mu^{(k)}) \to$ optimal primal/dual solution.