OR 635: Interior Point Methods for Math Progg. Fall 2003.

Comments on Final Exam.

1. a) You all did well on this, with the exception of showing that the solution to $(WBP)_\mu$ was unique. Note that, because some s_j^*'s are zero, the objective fuinction is *not* strictly convex in x. It *is* strictly convex as a function of the x_j's, $j \in N := \{i \in \{1, 2, \ldots, n\} : s_i^* > 0\}$. We can conclude that the values of such x_j's in any optimal solution are unique. But then, with $B := \{i \in \{1, 2, \ldots, n\} : s_i^* = 0\}$, we have $A_B x_B + A_N x_N = b$ for any feasible solution, so if x_N is unique, so is x_B, since you were given that A_B had full column rank.

I misled you slightly: the given solution only works for an open interval of μ's containing (0,1], since for large μ some components of $x(\mu, w)$ may be negative. Sorry.

b) You were mostly good on this, but your arguments were a bit sloppy. Maybe the easiest approach is to write down the optimality conditions, argue that these are necessary and sufficient and define a unique solution, hence show that x is on the path corresponding to $\mu = 1$, and use the implicit function theorem to show that the solution varies smoothly with μ. Then you can differentiate the conditions to get a linear system that the tangent must satisfy. Comparing this system to that defining the primal-dual affine-scaling direction, and noting that the solutions to these systems are unique, completes the proof.

2. You all did well on this, but simple crude estimates suffice to establish parts (a) and (b).

3. a) You were fine on this, but a little more precision on which conditions were necessary and which sufficient and why would have been welcome.

b) Consider any μ and $(x(\mu), y(\mu), s(\mu))$. Then $x(\mu)$ minimizes $(1/\mu)\langle c, x\rangle + F(x)$ over $\mathcal{F}^0(P)$, so satisfies

$$\frac{1}{\mu}c + F'(x) - A^* y = 0$$

for some y. Now $\mu\rho > \mu\nu = \langle c, x(\mu)\rangle - \langle b, y(\mu)\rangle$, so $\zeta := \langle c, x(\mu)\rangle - \mu\rho < \langle c, x(\mu)\rangle - \mu\nu = \langle b, y(\mu)\rangle \leq \zeta_*$, and then $\mu = (\langle c, x(\mu)\rangle - \zeta)/\rho$.

It follows that $x(\mu)$ satisfies

$$\frac{\rho}{\langle c, x\rangle - \zeta}c + F'(x) - A^* y = 0,$$

so it is a stationary point of $\phi_\rho(\cdot, \zeta)$ over $\mathcal{F}^0(P)$, i.e., it satisfies the necessary optimality condition. Indeed, by part (a), *any* point that satisfies the necessary condition must be $x(\mu)$ by part (a). So it suffices to show that $\phi_\rho(\cdot, \zeta)$ *has* a minimizer over $\mathcal{F}^0(P)$.

From the definition of ζ above, for any $x \in \mathcal{F}^0(P)$, $\langle c, x\rangle - \zeta > \langle c, x\rangle - \langle b, y(\mu)\rangle = \langle s(\mu), x\rangle \geq 0$. Thus for any such x,

$$\phi_\rho(x, \zeta) > \rho \ln\langle s(\mu), x\rangle + F(x) =: \psi(x).$$

Now note that $\bar{K} := \{x \in K : \langle s(\mu), x\rangle = \mu\nu\}$ is compact, so $\psi(x)$ (which is just $\rho\ln(\mu\nu) + F(x)$ there) attains its minimum, say $\underline{\psi}$, there by standard arguments.

Next note that, for any $x \in \mathcal{F}^0(P)$ with $\tau := \langle s(\mu), x\rangle/(\mu\nu) \geq 1$, we have

$$\psi(x) = \psi(x/\tau) + \rho\ln\tau - \nu\ln\tau \geq \underline{\psi} + (\rho - \nu)\ln\tau.$$

Thus for $x \in \mathcal{F}^0(P)$ with $\langle s(\mu), x\rangle \geq \bar{\tau}\mu\nu$, with $\bar{\tau}$ such that $(\rho - \nu)\ln\bar{\tau} = \phi_\rho(x(\mu), \zeta) - \underline{\psi}$, we find

$$\phi_\rho(x, \zeta) \geq \psi(x) \geq \underline{\psi} + (\rho - \nu)\ln\bar{\tau} = \phi_\rho(x(\mu), \zeta),$$

so we can confine our search for a minimizer of $\phi_\rho(\cdot,\zeta)$ to $\{x \in \mathcal{F}^0(P) : \langle s(\mu), x\rangle \leq \bar{\tau}\mu\nu\}$, which is bounded. Now standard arguments prove that a minimizer exists.

c) There was a little hint here in that parts (a) and (b) stated that $\rho > \nu$, but (c) doesn't. Indeed, if $\rho > \nu$, we showed that $\Phi_\nu(x,s)$ is unbounded below, so has no minimizers. This shows that an approach using parts (a) and (b) won't work, and that we need to have $\rho = \nu$. We know that $\Phi_\nu(x,s) \geq \nu \ln \nu - \nu$. Also, if $s + \mu F'(x) = 0$ for some $\mu > 0$, then $F'_*(s) = F_*(-\mu F'(x)) = -F(x) - \nu - \nu \ln \mu$, so that

$$\Phi_\nu(x,s) = \nu \ln(\mu\nu) + F(x) - F(x) - \nu - \nu \ln \mu = \nu \ln \nu - \nu,$$

so any such point minimizes $\Phi_\nu(x,s)$. Moreover, the proof of the lower bound shows that only such points minimize it. Hence,

the minimizers of Φ_ν over $\mathcal{F}^0(P) \times \mathcal{F}^0(D)$ are exactly the points on the primal-dual central path (omitting their y-components).

4. a) You all argued successfully that any optimal solution satisfied the general linear constraint at equality. But it would have been good to also argue that any optimal solution had all x_i's positive, so when you replace the objective by minimimizing $-\ln(x)$, there would be a feasible point with finite objective value. I believe I stated sometime that the minimizer of $F(x)$ over its domain (if it exists) is the analytic center of the domain, but if not, I should have!

Since you now have a convex objective and linear constraints, the KKT conditions (or just Lagrange conditions) are necessary and sufficient, whence you get the solution $x_1 = 24, /, x_2 = x_3 = 12$.

b) You all got something that was a SOCP, but often more complicated than necessary. Note that $x_1x_2 \geq 1$, $x_1, x_2 \geq 0$ iff $(x_1+x_2)^2 - (x_1-x_2)^2 \geq 4$, $x_1, x_2 \geq 0$, and this holds iff $((x_1+x_2);(x_1-x_2);2) \in S^3_2$. (The nonnegativities follow from $x_1 + x_2 > x_1 - x_2$ and $x_1 + x_2 \geq -(x_1 - x_2)$.) So the optimization problem can be written in dual form as

$$\max\{-2(x_1+x_2) : ((x_1+x_2);(x_1-x_2);2) \in S^3_2\}.$$

Noting that the mapping from x to $(\xi_1;\xi_2)$ with $\xi_1 = x_1 + x_2$, $\xi_1 = x_1 + x_2$ is one-to-one and onto, we can also write it in primal form as

$$\min\{2\xi_1 : \xi_3 = 2, \xi \in S^3_2\}.$$

The optimal solution is then easy to obtain. From $((x_1+x_2);(x_1-x_2);2) \in S^3_2$ we see that $2(x_1+x_2) \geq 4$, and that equality holds only if $x_1 - x_2 = 0$. But this is possible by setting $x_1 = x_2 = 1$. (If you don't know about it: Flatland : a romance of many dimensions / Edwin A. Abbott ; illustrations by the author A. Square ; foreword by Isaac Asimov. Sphereland : a fantasy about curved spaces and an expanding universe / by Dionys Burger; translated from the Dutch by Cornelie J. Rheinboldt ; foreword by Isaac Asimov. QA699 .A13 1994.)