Fairly practical IP methods - example ①

work with a linear program in standard form

$$\min \; c^T x \quad \text{st} \quad Ax = b, \quad x \geq 0$$

[Notice ineq reversed sign from above]
that's life

dual is

$$\max \; b^T \lambda \quad \text{st} \quad A^T \lambda + s = c, \; s \geq 0$$

[check this!]

KKT are :

$$A^T \lambda + s = c$$
$$Ax = b$$
$$x_i s_i = 0$$
$$x \geq 0 \quad s \geq 0$$

General approach

- start w/ x, λ, s, <u>feasible</u>
- update w root finding, subst.

  $$x_i s_i = (\text{something } +ve)$$

  for

  $$x_i s_i = 0$$
- choose steps, etc st $x_i s_i$ gets smaller AND x, λ, s remains feasible

Equation set up for root finding

Dual residual

$$\cancel{A^T s =} \quad A^T \lambda + s - c = 0$$

Centering residual

$$\mathrm{diag}(s)\, x - \left(\substack{\text{something} \\ +ve}\right) = 0$$

Primal residual

$$Ax - b = 0$$

Root finding setup:

- at $x^{(i)}, \lambda^{(i)}, s^{(i)}$, and seek $\Delta x, \Delta\lambda, \Delta s$, so that we are at root.

- This yields:

$$\begin{bmatrix} 0 & A^T & I \\ \mathrm{diag}(s^{(i)}) & 0 & \mathrm{diag}(x^{(i)}) \\ A & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} -\left[A^T\lambda^{(i)} + s^{(i)} - c\right] \\ (\text{+ve thing}) - \mathrm{diag}(s^{(i)})x^{(i)} \\ -\left[A x^{(i)} - b\right] \end{bmatrix}$$

- look at first – interpret as measure of dual feasibility (Basically, choose $\Delta\lambda$, Δs to become dual feasible)

  third – "primal" (" Δx " primal)

Useful concept

· The duality measure

$$\mu = \frac{1}{N} \sum x_i s_i$$

\# of vars

· Notice analogy with $\left\{ \begin{array}{l} \text{central path} \\ \text{log barrier} \end{array} \right\}$ case.

where $\mu_c = \frac{t}{N}$ was duality gap.

· if μ is small, and x, λ, s are all feasible, we have to be close to sol'n

Now consider the (+ve thing)

- <u>cases</u>

0 — choose update to solve KKT.
- takes us toward constraints
- known as "affine scaling dir"
- usually, not much progress

$1/t$ — choose update to go on central path.
- implicitly, t is increasing.

$\sigma^{(i)} \mu^{(i)}$ — $\sigma^{(i)} > 0$, but $\sigma < 1$
- ie try to make duality measure better.
- this tends to allow better steps

we do

(5)

$$x^{(i+1)}, \lambda^{(i+1)}, s^{(i+1)} = x^{(i)}, \lambda^{(i)}, s^{(i)} + \alpha(\Delta x, \Delta\lambda, \Delta s)$$

issues:

- choose $\sigma^{(i)}$, α.
- α must be such that $x^{(i+1)} > 0$, $s^{(i+1)} > 0$

Q: What α should we accept?

A: ~~$\alpha \in$ some neighbo~~

$x^{(i+1)}, \lambda^{(i+1)}, s^{(i+1)}$ in some neighborhood of the central path

$$\mathcal{N}_{-\infty}(\gamma) = \{ x, \lambda, s \in \mathcal{F} \mid x_i s_i \geq \gamma\mu \text{ for each } i \}$$

γ typically $1e^{-3}$

these are points s.t. each comp constraint is ~~most~~ bigger than $\gamma \cdot$ [duality measure]

Typically, choose

$$\alpha \in [0,1]$$

largest, st $x^{(i+1)}, \lambda^{(i+1)}, s^{(i+1)} \in N_{-\infty}(\gamma)$.

- This allows steps to come quite close to some constraints.

<u>Even better</u>:

- take <u>different</u> steps in primal x vars and dual λ, s.

<u>write</u>

$$\alpha_{max}^{pri,i} = \min_{u:\, \Delta x_u < 0} \frac{-x_u^{(i)}}{\Delta x_u^{(i)}}$$

this is the largest step we can take in primal w/o violating constraints

Similarly:

$$\alpha_{max}^{dual,i} = \min_{u:\, \Delta s_u^{(i)} < 0} -\frac{s_u^{(i)}}{\Delta s_u}.$$

Often use:

$$\alpha^{primal,i} = \min\left(1, \eta^{(i)} \cdot \alpha_{max}^{primal,i}\right)$$

$$\alpha^{dual,i} = \min\left(1, \eta^{(i)} \cdot \alpha_{max}^{dual,i}\right)$$

where $\eta^{(i)} \in [0.9, 1)$ and $\eta^{(i)} \longrightarrow 1$

Notice that these step's must drive down infeasibility (N+W, 2e, p408)

Corrector steps

Imagine we solve

$$\begin{bmatrix} 0 & A^T & I \\ d(s) & 0 & d(x) \\ A & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} -[A^T\lambda + s - c] \\ -d(s)x \\ -[Ax - b] \end{bmatrix}$$

and take a full step in the resulting dirn.

Then

$$(x_u^{(i)} + \Delta x_u)(s_u^{(i)} + \Delta s_u)$$

$$= x_u^{(i)} s_u^{(i)} + x_u^{(i)} \Delta s + s_u^{(i)} \Delta x + \Delta x \Delta s$$

$$= \Delta x \Delta s$$

which isn't zero

we could fix this by taking <u>another</u> step, given by

$$\begin{bmatrix} 0 & A^T & I \\ d(s) & 0 & d(x) \\ A & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x^* \\ \Delta \lambda^* \\ \Delta s^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -d(s).x \end{bmatrix}$$

this is a <u>corrector step</u>

in many cases,

$$(\Delta x, \Delta \lambda, \Delta s) + (\Delta x^*, \Delta \lambda^*, \Delta s^*)$$

<u>is a better step</u>.