Interior Point Algos :

Suppose $P = \{x : Ax \geq b\}$ ← say polytope. want to min ~~...~~ $= c^Tx$ st $x \in P$.

Constrained problem. Let's throw them away; replace by a "barrier"

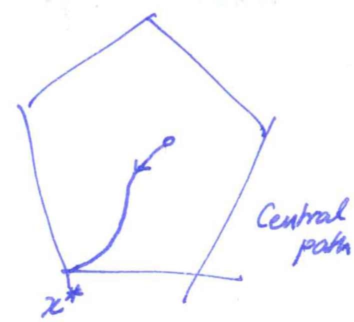
$$\min f_\eta(x) = c^Tx + \eta \cdot \sum_{i=1}^{m} \ln\left(\frac{1}{a_i x - b_i}\right)$$

So as $a_i x \to b_i$ the function tends to infty. (Hence the function is making a barrier as we reach the boundary of P).

. Hence as $\mu \neq 0$, we need to assume that the polytope P has a non-empty interior.

Now the minimizer of $f_\eta(x)$ is called the "analytic center". Note it depends on η. But also on the representation of P. E.g. if you double the constraint # i then the function f_η and the associated analytic center changes.

Now consider the path as $\eta \to 0$. In that case this ~~path~~ x^*_η analytic center goes to the actual minimizer.



Algorithms that (try to) follow this path are called path-following methods. Idea: if you have a point x_η (close to x^*_η) then getting a point $x_{\eta'}$ (close to $x^*_{\eta'}$) is not that difficult. So keep moving "close to" the central path until very close to x^*, and then round much like in ellipsoid

---

OK. Let's consider an LP in equational form for the rest of the discussion.

$\min c^Tx$		$\max b^Ty$
$Ax = b$	$\iff$	$A^Ty \leq c$
$x \geq 0$.		
(primal)		(dual)

Also assume that both primal and dual are strictly feasible. Both have solutions in their interiors.

technical requirement

To solve this again add the barrier to get

$$\min f_\eta(x) = c^T x - \eta \sum_{i=1}^{n} \ln x_i$$

$$Ax = b.$$

advantage of equational form.

How to find the minimizer? Method of Lagrange Multipliers.

"Lagrangify" the constraints

$$\min \quad f_\eta(x) - \sum_{i=1}^{m} y_i (a_i x - b_i) \quad = \quad f_\eta(x) - y^T(Ax-b).$$

At the minimizer we should have gradients $=0$.

gradient wrt $\bar{x}$: $\nabla f_\eta(x) - y^T A = 0$

wrt $\bar{y}$: $Ax - b = 0$.

What is $\nabla f_\eta(x) = c^T - \eta\left(\frac{1}{x_1}, \frac{1}{x_2} \cdots, \frac{1}{x_n}\right) \Rightarrow$ $Ax = b$

$c^T - \eta(1/\bar{x})^T = y^T A$

notation for $(1/x_1 \cdots 1/x_n)$

$$\Rightarrow \boxed{\begin{aligned} &Ax = b \\ &A^T y + S = c \\ &\bar{S} = \eta(1/\bar{x}) \quad \text{or} \quad \bar{s}_i x_i = \eta \ \forall i \end{aligned}}$$

"slack"

Two comments:

① Lagrange multipliers show necessity of existence of y, s such that these cond'ns are satisfied at the minimum. Need to show they are sufficient to capture the minimum, need to use the strict feasibility for that. See [Matousek Gaertner] for details.

② Sps $\eta = 0$ then this is exactly complementary slackness!

$$c^T x = (y^T A x + s x) = y^T A x = y^T b \Rightarrow \text{primal} = \text{dual}!$$

Good. We have the system $Ax = b \quad A^T y + s = c$] maintain exactly

$s \circ x = \eta \mathbb{1}$ maintain approximately.

In particular make sure that

$$\| s \circ x - \eta \mathbb{1} \|_2 \le \theta \cdot \eta \quad \text{for } \boxed{\theta = 0.4.} \qquad (*)$$

Algo: (1) Start with $s_0, x_0 > 0$, y_0 st $Ax_0 = b$, $A^T y_0 + s_0 = c$

How? see books...

and $(*)$ is satisfied for η_0. Define $\eta_0 = \frac{\langle x_0, s_0 \rangle}{n}$

(2) at each time t, find soln $(\Delta x, \Delta y, \Delta s)$ to

solvable by "necessary" part of Lagrange mult argument.

$$A \Delta x = 0$$
$$A^T \Delta y + \Delta s = 0$$
$$(s \circ \Delta x) + (x \circ \Delta s) = -(x_{t-1} \circ s_{t-1}) + \eta_t \cdot \mathbb{1}$$

where $\eta_t = \sigma \cdot \eta_{t-1}$ $\boxed{\sigma = \left(1 - \frac{\theta}{\sqrt{n}}\right)}$

↑ intuition: drop $(\Delta x \circ \Delta s)$ quadratic term.

set $x_t \leftarrow x_{t-1} + \Delta x$ etc.

Claims: (1) $\forall t$, s_t, x_t, y_t satisfy $Ax = b \quad A^T y + s = c$. (By construction)

To prove! → (2) $\forall t$ s_t, x_t satisfy $\| x_t \circ s_t - \eta_t \mathbb{1} \| \le \theta \eta_t \quad (*)$

(3) $\eta_t = \sigma^t \cdot \eta_0 \le \sigma^t \cdot 2^{\Theta(L)}$ where $L = \langle A \rangle + \langle b \rangle + \langle c \rangle$

$\Rightarrow$ after $t = O(\sqrt{n} L)$ rounds $\eta_t \le 2^{-\Theta(L)}$.

(4) $\langle \Delta x, \Delta s \rangle = 0 \quad \forall$ times.

Pf: Indeed $\Delta s = -A^T \Delta y \Rightarrow \langle \Delta x, \Delta s \rangle = -\Delta x^T A^T \Delta y = 0.$ ∎

(5) $\eta_t = \frac{\langle x_t, s_t \rangle}{n}$

Pf: $\langle x^t, s^t \rangle = \langle \mathbb{1}, x^t \circ s^t \rangle = \langle \mathbb{1}, \eta_t \mathbb{1} - \Delta x_t \circ \Delta s_t \rangle = \eta_t \cdot n - 0$ (by 4)

∎

(6) After $T = O(\sqrt{n}L)$ steps almost optimal sol^{n}s.

$$\eta_t = \frac{1}{n}\langle x_t, s_t\rangle = \frac{1}{n}\langle x_t, -A^T y_t + c\rangle = \frac{1}{n}\langle c^T x_t - b^T y_t\rangle$$
$$= \frac{1}{n}(\text{duality gap})$$

$\Rightarrow x_t, y_t$ are both nearly optimal upto $2^{-O(L)}$ error.

Finally, remains to prove #2 above. Pf not particularly edifying, but here it is.

Pf: $\| x_t \circ s_t - \mathbb{1}.\eta_t \|_2 = \| \Delta x \circ \Delta s \|_2$ by construction. And while the inner product $\langle \Delta x \circ \Delta s \rangle = 0$ no control per se on per-entry.

$x = x_{t-1}$
$s = s_{t-1}$

Let $D = \mathrm{diag}(\sqrt{x_i/s_i})$.

then
$$\|\Delta x \circ \Delta s\| = \|(D^{-1}.\Delta x) \circ (D.\Delta s)\|$$
$$\leq 2^{-3/2} . \| D^{-1}\Delta x + D\Delta s \|^2$$
$$= 2^{-3/2} \| \sqrt{\tfrac{1}{x \circ s}} (s\Delta x + x\Delta s) \|^2$$
$$= 2^{-3/2} \| \sqrt{\tfrac{1}{x} \circ \tfrac{1}{s}} (- x_{t-1} \circ s_{t-1} + \eta_t \mathbb{1}) \|^2$$
$$= 2^{-3/2} \sum_i \frac{(-x_i s_i + \eta_t)^2}{x_i s_i} \leq 2^{-3/2} \frac{\| x_{t-1} \circ s_{t-1} - \eta_t \mathbb{1} \|^2}{\min x_i s_i}$$

$\|a \circ b\| \leq 2^{3/2} \|a+b\|^2$

but $x_i s_i \geq 1 - \theta . \eta_{t-1}$ (by induction)

and $\| x_{t-1} \circ s_{t-1} - \eta_t \mathbb{1} \|^2$
$$= \| (x_{t-1} \circ s_{t-1} - \eta_{t-1}.\mathbb{1}) + (1-\sigma)\eta_{t-1}.\mathbb{1} \|^2$$
$$= \| x_{t-1} \circ s_{t-1} - \eta_{t-1}\mathbb{1} \|^2 + (1-\sigma)^2 \| \eta_{t-1} \mathbb{1} \|^2 + 0$$
$\rightarrow$ since $(x_{t-1} \circ s_{t-1} - \eta_t \mathbb{1}) \perp \mathbb{1}$. (by 5)
$$\leq \theta^2 \eta_{t-1}^2 + (1-\sigma)^2 . \eta_{t-1}^2 . n^2$$

If $\sigma = 1 - \theta/\sqrt{n} \Rightarrow 2\theta^2 \eta_{t-1}^2$. And the $2^{-3/2}$ factor can give us $\leq \theta^2 \eta_t^2$. (strict)

∎

Let's say the argument again.

Start with $x_0, s_0 > 0$. define $\eta_0 = \frac{\langle x_0, s_0\rangle}{n}$

each time x_t, s_t obtained by adding in $\Delta x, \Delta s$. Maintain $\eta_t = \frac{\langle x_t, s_t\rangle}{n}$

by fact 5.

Finally get x_t, y_t to ~~be~~ have $c^T x_t - b^T y_t \le \eta_t$ and stop.

<u>What about feasibility?</u>

Clearly $Ax = b$, $A^T y + s = c$ satisfied. But $x, s \geq 0$?

Infact maintain $x, s > 0$.

Start with such a soln. $x_0, s_0 > 0$.

Now claim that preserved at time t.

(a) if some $x_i s_i \le 0$ then ℓ_2-distance would give no η_t just from that coord, and hence violate (*). In fact $x_i s_i < \eta_t \times 0.6$ would do it.

(b) what if both x_i and $s_i < 0$ but $x_i s_i > 0$?

Can actually ~~tune~~ the argument to show that not just $(x_t \circ s_t)$ satisfies (*) but all points of form $(x_{t-1} + \alpha \Delta_t) \circ (s_t + \alpha \Delta_t)$ for $\alpha \in [0, 1]$ satisfy (*). So all is feasible.
