Primal Dual methods

when we followed a path, we solved

$$\min_{x} \quad t f(x) + \varphi(x)$$
$$\text{st} \quad Ax = b$$

during process of solving, we found x^*, λ_i^*, ν^*
$\left(\lambda_i^* = \dfrac{-1}{t\, g_i(x^*)} \right.$ and $\left. \nu^* = \frac{1}{t} \nu \right)$

Original problem was

$$\min \quad f(x)$$
$$\text{st} \quad g_i(x) \leq 0$$
$$Ax = b$$

Analogy

Properties of x^*, λ_i^*, ν^*	KKT for original problem
$Ax^* - b = 0$	$Ax - b = 0$
$g_i(x) < 0$	$g_i(x) \leq 0$
$\lambda_i^* > 0$	$\lambda_i \geq 0$
$\nabla f + \sum_i \lambda_i^* \nabla g_i + \nu^{*T} A = 0$	$\nabla f + \sum_i \lambda_i \nabla g_i + \nu^T A = 0$
$-\lambda_i^* g_i = \frac{1}{t}$	$\lambda_i g_i = 0$

Notice: (I) we have feasible x^*, λ^*, ν^* and by increasing t we are pushing towards complementarity

IDEA: do this explicitly.

Strategy:

for given t, find x^*, λ_i^*, ν^* to solve

$$Ax^* - b = 0$$

$$\nabla f + \sum_i \lambda_i^* \nabla g_i + \nu^{*T} A = 0$$

$$-\lambda_i^* g_i(x^*) - \frac{1}{t} = 0$$

and do so in a way that $\lambda_i^* > 0$, $g_i(x^*) < 0$

even if f, g_i are linear, this system isnt

$$-\lambda_i^* g_i(x^*) - \frac{1}{t} = 0$$

!

· See this as root-finding

The central path

points $x^*(t)$, $\lambda_i^*(t)$, $\nu^*(t)$, $\tau = \frac{1}{t}$

That (A) satisfy eqns of P. 3

(B) $\lambda_i^* > 0$, $g_i(x^*) < 0$

This is a curve. as $\tau \left(= \frac{1}{t}\right) \to 0$, tends to soln

Procedure: find a sequence of points on this path, with decreasing τ (increasing t).

Root-finding:

$$G(u^*) = 0$$

u^*: d-dim var

G: d-dim vector of fn's

write $J_{G,u} = \begin{bmatrix} \frac{\partial g_1}{\partial u_1} & \frac{\partial g_1}{\partial u_2} & \cdots \\ \frac{\partial g_2}{\partial u_1} & \ddots & \\ \vdots & & \ddots \end{bmatrix}$ (Jacobian; Derivative; matrix of partial derivatives)

then assume we have $u^{(i)}$ an estimate

$$G(u^{(i)} + \Delta u) \approx G(u^{(i)}) + J_G \, \Delta u$$

so we want

$$J_G \, \Delta u = -G(u^{(i)})$$

Look at system more closely

$$\nabla f + \sum_i \lambda_i \nabla g_i + \nu^T A$$

this is the dual residual. if it is 0, then λ, ν are dual feasible.

$$\underbrace{-\operatorname{diag}(\lambda)\, g - \frac{1}{t}\mathbf{1}}_{\text{vector of eqns, 1 per } g_i}$$

centering residual (is the λ, x close to comp.?)

$$Ax - b$$

primal residual (is x feasible)

these are written

r_d

r_c

r_p

$\tau = (1/t)$ appears only in centering residual. ⑥

so our Newton system looks like

$$\begin{bmatrix} \text{Big Matx} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta\nu \end{bmatrix} = \begin{bmatrix} \text{no } \tau \\ -\lambda_i^k g_i^k - \tau \\ \text{no } \tau \end{bmatrix}$$

one procedure *might be*

- choose small τ
- compute $\begin{matrix} \Delta x \\ \Delta\lambda \\ \Delta\nu \end{matrix}$
- $\begin{matrix} x^{k+1} \\ \lambda^{k+1} \\ \nu^{k+1} \end{matrix} = \begin{matrix} x^k \\ \lambda^k \\ \nu^k \end{matrix} + \alpha \begin{matrix} \Delta x \\ \Delta\lambda \\ \Delta\nu \end{matrix}$

where we get α from search.

Q: how to choose τ st:

- search is easy
- we get good progress

A measure of progress.

$$\frac{\sum_i -\lambda_i^k g_i(x^k)}{\#\ \text{constraints}} = \mu$$

is the duality measure

this is an approx est of the primal-dual gap at x^k, λ_i^k