Primal-dual interior point method (Mattingley & Boyd)

minimize $\frac{1}{2}x^TQx+q^Tx$ $\quad Q\in S^n_+,\ q\in\mathbb{R}^n$

subject to $Gx\le h,\ Ax=b$ $\quad G\in\mathbb{R}^{p\times n},\ h\in\mathbb{R}^p,\ A\in\mathbb{R}^{m\times n},\ b\in\mathbb{R}^m$

$Gx\le h$

$Gx-h\le 0$

$\downarrow$ Introduce slack variables

symmetrizing a matrix:

$$M=\underbrace{\frac{M+M^T}{2}}_{\text{symmetric part}}+\frac{M-M^T}{2}$$

minimize $\frac{1}{2}x^TQx+q^Tx$

subject to $Gx+s=h,\ Ax=b,\ s\ge 0$

$\downarrow$ KKT conditions

z: inequality dual

y: equality dual

For minimize $f(x)$

$h_i(x)\le 0$

$l_j(x)=0$

$\nabla f_0(x)+\cancel{G^Tz}+\sum z_i\,\partial h_i(x)+\sum y_j\,\partial l_j$

$Gx+s=h,\ Ax=b,\ s\ge 0$ primal feasibility

$z\ge 0,\ (z\in\mathbb{R}^p)$ dual feasibility

$Qx+q+G^Tz+A^Ty=0$ stationarity

$z_i s_i=0,\ i=1,\dots,p$ complementary slackness

$h(x)=Gx-h\le 0$

$\partial h(x)=G$

$l(x)=Ax-b=0$

$\partial l=A$

Initialization

Solve for optimality conditions of both primal and dual problems (of related problem) — read CVXopt paper

Analytically

$$\begin{bmatrix}Q & G^T & A^T\\ G & -I & 0\\ A & 0 & 0\end{bmatrix}\begin{bmatrix}x\\ z\\ y\end{bmatrix}=\begin{bmatrix}-q\\ h\\ b\end{bmatrix}$$

TODO work this out

Gives starting point $(x^{(0)}, s^{(0)}, z^{(0)}, y^{(0)})$

~~KKT matrix~~

$Kl=r$ — want efficient way to do $l=K^{-1}r$

Main iteration

1. Evaluate stopping criteria (residuals, duality gap)
2. Compute affine scaling directions $\Delta x^{aff}, \Delta s^{aff}, \Delta z^{aff}, \Delta y^{aff}$
3. Compute ~~centering~~ centering-plus-corrector directions $\Delta x^{cc}, \Delta s^{cc}, \Delta z^{cc}, \Delta y^{cc}$
4. Update primal and dual variables

both use this

$$\begin{array}{c}n\\ n\\ n_{ineq}\\ n_{eq}\end{array}\begin{bmatrix}Q & 0 & G^T & A^T\\ 0 & Z & S & 0\\ G & I & 0 & 0\\ A & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\Delta x\\ \Delta s\\ \Delta z\\ \Delta y\end{bmatrix}=\ \text{aff},\ \text{cc}$$

symmetrize this

~~make KKT~~ Lots of structure (see 5.3 for explanation)

symmetric given conditions

$$PKP^T=LDL^T$$

permutation matrix; diagonal; unit lower-triangular

Amos, Kolter use LU factorization. Doesn't matter for us

We can even just do inv(K) at first

Two more tricks:

- Dynamic regularization — to avoid divide-by-zero errors
- Choosing permutation P that leads to sparse L

Regularize K to ensure factorization always exists, numerically stable.

$$\tilde{K}=K+\left[\begin{array}{c|c}\varepsilon I & 0\\ \hline 0 & -\varepsilon I\end{array}\right]$$

$$\boxed{l}=\tilde{K}^{-1}r=P^TL^{-T}D^{-1}L^{-1}Pr$$

solutions to perturbed system

~~con~~, just use this: $\tilde{K}l=r$ $\longrightarrow$ Iterative refinement to move towards $Kl=r$

K depends on z and s

$Z=\mathrm{diag}(z)$, $S=\mathrm{diag}(s)$

OptNet

$$\underset{z}{\text{minimize}}\ \frac{1}{2}z^TQz+q^Tz \longrightarrow L(z,\nu,\lambda)=\frac{1}{2}z^TQz+q^Tz+\nu^T(Az-b)+\lambda^T(Gz-h)$$

lagrangian

subject to $Az=b,\ Gz\le h$

- KKT conditions

$Qz^*+q+A^T\nu^*+G^T\lambda^*=0$ (stationarity)

$Az^*-b=0$ (primal feasibility)

$D(\lambda^*)(Gz^*-h)=0$ (complementary slackness)

they ignore this $\rightarrow$ $\lambda^*\ge 0$ (dual feasibility)

Derivatives

$$dQz^*+Qdz+dq+dA^T\nu^*+A^Td\nu+dG^T\lambda^*+G^Td\lambda=0$$

$$dAz^*+Adz-db=0$$

$$D(Gz^*-h)\,d\lambda+D(\lambda^*)(dGz^*+Gdz-dh)=0$$

Seems like rules for matrix differentials are pretty much same as scalars

matrices form

$$\underbrace{\begin{bmatrix} Q & G^T & A^T \\ D(\lambda^*)G & D(Gz^*-h) & 0 \\ A & 0 & 0 \end{bmatrix}}_{\text{call this } H}\begin{bmatrix} dz \\ d\lambda \\ d\nu \end{bmatrix}=\begin{bmatrix} -dQz^*-dq-dG^T\lambda^*-dA^T\nu^* \\ -D(\lambda^*)dGz^*+D(\lambda^*)dh \\ -dAz^*+db \end{bmatrix}$$

$$\begin{bmatrix} d_z \\ d_\lambda \\ d\nu \end{bmatrix}=H^{-1}\begin{bmatrix} \left(\frac{\partial L}{\partial z^*}\right)^T \\ 0 \\ 0 \end{bmatrix}$$

This comes from chain rule $\frac{\partial L}{\partial z^*}\frac{\partial z^*}{\partial b}$

This gives us gradients

Not fastway $\rightarrow$ Found by setting $db=I$ and setting all other differential terms in right side to zero, then solving equation.

If we want $\frac{dz^*}{dh}$:

$$\begin{bmatrix} dz \\ d\lambda \\ d\nu \end{bmatrix}=H^{-1}\begin{bmatrix} 0 \\ D(\lambda^*)dh \\ 0 \end{bmatrix} \quad dh=I$$