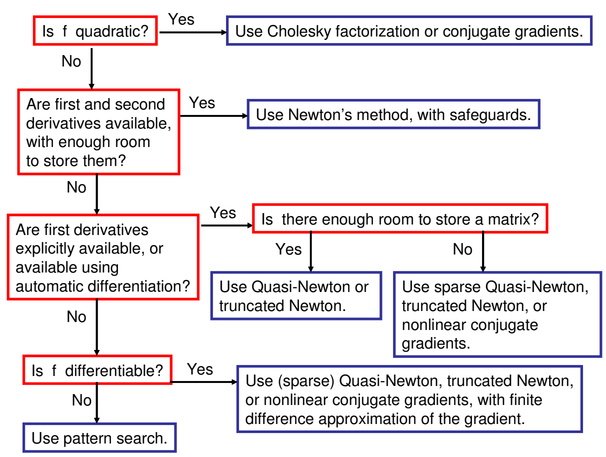
SOLUTION TO

AMSC 607 /CMSC 764 Final Exam , Fall 2010

1.



2. The Lagrangian is

$$L(\boldsymbol{x}, \boldsymbol{y}, z, \boldsymbol{w}) = e^{(x_1-5)}+6x_2^4+x_1x_2+x_3^2+x_4^2+3x_2+5+\boldsymbol{y}^T(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})-z(1-x_1^2-x_2^2)-\boldsymbol{w}^T\boldsymbol{x}.$$

To get the first order necessary conditions, we take the partial derivatives and add sign conditions and complementarity for Lagrange multipliers corresponding to inequality constraints:

$$
\begin{array}{rcl}
\begin{bmatrix} e^{(x_1-5)}+x_2 \\ 24x_2^3+x_1+3 \\ 2x_3 \\ 2x_4 \end{bmatrix} + \boldsymbol{A}^T\boldsymbol{y} - z\begin{bmatrix} -2x_1 \\ -2x_2 \\ 0 \\ 0 \end{bmatrix} - \boldsymbol{w} & = & \boldsymbol{0}, \\
\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} & = & \boldsymbol{0}, \\
1 - x_1^2 - x_2^2 & \geq & 0, \\
\boldsymbol{x} & \geq & \boldsymbol{0}, \\
z & \geq & 0, \\
\boldsymbol{w} & \geq & \boldsymbol{0}, \\
z(1 - x_1^2 - x_2^2) & = & 0, \\
\boldsymbol{w}^T\boldsymbol{x} & = & 0.
\end{array}
$$

3. We can solve the 1st equation for $\boldsymbol{x}$ and the 3rd for $\boldsymbol{z}$:

$$
\begin{array}{rcl}
\boldsymbol{x} & = & \boldsymbol{D}^{-1}(\boldsymbol{a} - \boldsymbol{E}\boldsymbol{z}), \\
\boldsymbol{z} & = & \boldsymbol{F}^{-1}(\boldsymbol{c} - \boldsymbol{B}\boldsymbol{y}).
\end{array}
$$

Substituting these expressions into the 2nd equation gives

$$\boldsymbol{A}\boldsymbol{D}^{-1}(\boldsymbol{a} - \boldsymbol{E}\boldsymbol{z}) + \boldsymbol{Q}\boldsymbol{y} = \boldsymbol{b},$$

so

$$\boldsymbol{A}\boldsymbol{D}^{-1}(\boldsymbol{a} - \boldsymbol{E}\boldsymbol{F}^{-1}(\boldsymbol{c} - \boldsymbol{B}\boldsymbol{y})) + \boldsymbol{Q}\boldsymbol{y} = \boldsymbol{b}.$$

Therefore, we can determine $\boldsymbol{y}$ by solving the linear system

$$(\boldsymbol{Q} + \boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{E}\boldsymbol{F}^{-1}\boldsymbol{B})\boldsymbol{y} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{D}^{-1}(\boldsymbol{a} - \boldsymbol{E}\boldsymbol{F}^{-1}\boldsymbol{c}).$$

We need extra information to guarantee that this linear system has a solution and that the solution is unique, but that was not part of the problem.

4. Nocedal and Wright, plus the class notes, are a good reference for this. (There may be a few sign errors in this code, and a few typos in the right-hand sides.)

Input: $\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}$; tolerance ϵ; feasible starting point $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ close to the central path. *(So* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *and* $\mathbf{A}^T\mathbf{y} + \mathbf{z} = \mathbf{c}$*.)*

Define parameter $\eta = .99$ and $\boldsymbol{e} = [1, \ldots, 1]^T$ $(n \times 1)$.

while $|\boldsymbol{b}^T\boldsymbol{y} - \boldsymbol{c}^T\boldsymbol{x}| > \epsilon$

- *The predictor (affine) step:*

  Let $\boldsymbol{D}^{-1} = \boldsymbol{X}^{-1}\boldsymbol{Z}$ and find $\boldsymbol{\Delta y}^a$ from the normal equations

$$\boldsymbol{A}\boldsymbol{D}^2\boldsymbol{A}^T\boldsymbol{\Delta y}^a = -\boldsymbol{A}\boldsymbol{Z}^{-1}\boldsymbol{r},$$

where

$$\boldsymbol{r} = -\boldsymbol{X}\boldsymbol{z}.$$

This is solved using a Cholesky decomposition, which is reused in the corrector step below.

- Compute

  $$\boldsymbol{\Delta z}^a = -\boldsymbol{A}^T\boldsymbol{\Delta y}^a$$

  and

  $$\boldsymbol{\Delta x}^a = \boldsymbol{Z}^{-1}(\boldsymbol{r} - \boldsymbol{X}\boldsymbol{\Delta z}^a).$$

  (Exploit the diagonal structure of the matrices to keep the cost low.)

- Determine the steplengths α_p^a and α_d^a from

  $$\begin{aligned} \alpha_p^a &= \min(1, \min_{i:\boldsymbol{\Delta x}_i^a<0}(-x_i/\boldsymbol{\Delta x}_i^a)), \\ \alpha_d^a &= \min(1, \min_{i:\boldsymbol{\Delta z}_i^a<0}(-z_i/\boldsymbol{\Delta z}_i^a)). \end{aligned}$$

- Set

  $$\mu^a = (\boldsymbol{x} + \alpha_p^a\boldsymbol{\Delta x}^a)^T(\boldsymbol{z} + \alpha_d^a\boldsymbol{\Delta z}^a)/n,$$

  where n is the length of $\boldsymbol{x}$, and compute

  $$\sigma = \left(\frac{\mu^a}{\mu}\right)^3,$$

  where $\mu = \boldsymbol{x}^T\boldsymbol{z}/n$.

- *The centering/corrector step:* Find $\boldsymbol{\Delta y}^c$ from the normal equations

  $$\boldsymbol{A}\boldsymbol{D}^2\boldsymbol{A}^T\boldsymbol{\Delta y}^c = -\boldsymbol{A}\boldsymbol{Z}^{-1}\boldsymbol{r},$$

  where

  $$\boldsymbol{r} = \sigma\mu\boldsymbol{e} - \boldsymbol{X}\boldsymbol{z} - \boldsymbol{\Delta X}^a\boldsymbol{\Delta z}^a,$$

  using the previous Cholesky decomposition.

- Compute

  $$\boldsymbol{\Delta z}^c = -\boldsymbol{A}^T\boldsymbol{\Delta y}^c$$

  and

  $$\boldsymbol{\Delta x}^c = \boldsymbol{Z}^{-1}(\boldsymbol{r} - \boldsymbol{X}\boldsymbol{\Delta z}^c).$$

- Determine the step lengths α_p^c and α_d^c from

  $$\begin{aligned} \alpha_p^c &= \min(1, \eta \min_{i:\boldsymbol{\Delta x}_i^c<0}(-x_i/\boldsymbol{\Delta x}_i^c)), \\ \alpha_d^c &= \min(1, \eta \min_{i:\boldsymbol{\Delta z}_i^c<0}(-z_i/\boldsymbol{\Delta z}_i^c)). \end{aligned}$$

- Take a step:

$$\begin{aligned} \boldsymbol{x} &= \boldsymbol{x} + \alpha_p^c \boldsymbol{\Delta x}^c, \\ \boldsymbol{y} &= \boldsymbol{y} + \alpha_p^d \boldsymbol{\Delta y}^c, \\ \boldsymbol{z} &= \boldsymbol{z} + \alpha_p^d \boldsymbol{\Delta z}^c, \end{aligned}$$

end while

5a. A polynomial complexity bound assures us that the algorithm won't take a very large number of iterations (e.g., exponential in the size of the problem).

5b. The main ingredients include:

- Use of a self-concordant barrier function, the log function. This ensures that iterates do not leave the feasible region but that the barrier does not grow so fast near the boundary that Newton's method gets lost when close to the solution.
- Use of a continuation method. (He uses η as the parameter.) This ensures that we can start close enough to the solution to the next problem to be in the region of fast convergence for Newton's method.
- A quantitative theorem on convergence of Newton's method, bounding the size of the Newton step in terms of the size of the previous step. This, plus the fact that the error can be bounded by the size of the Newton step, gives us a measure of how near the solution we are.
- A careful choice of "close enough" (in the "Putting it all together" section) that enables us to say that if we start close to a point on the central path, and we take a Newton step, we will be close to a point on the central path that is measurably closer to the solution to our problem.
- A computable bound on how large η needs to be in order for us to be within ϵ of the optimal value of $\boldsymbol{c}^T\boldsymbol{x}$.

The conclusion is that the algorithm takes $O(\sqrt{n}\log(n/\epsilon))$ iterations.

5c. The work of *Fiacco* and McCormick on barrier methods for solving constrained optimization problems included most of the important ideas behind the very successful interior point methods (IPMs) used today, but in some sense they were *too_soon*, missing the very large problems that motivate the use of IPMs, the complexity proof, and the linear algebra knowledge that makes the algorithms work reliably.