Linear Programming

(5531)

Lecture 09 Interior Point Methods (II)

Prof. Sheng-I Chen

http://www.iem-omglab.nctu.edu.tw/OMG_LAB/

November 19, 2020

The Path-following Method
(Chapter 18)

The Path-following Method

- The **path-following method** belongs to a class of **interior-point methods**
- An alternative to the simplex method for solving LP
- Similar to the simplex method, the path-following method uses of iterative approach as the following:
  1. Find a initial solution
  2. Check if the current solution is good enough
  3. If not, repeat the previous steps to find a better solution

Karmarkar Algorithm Proves Its Worth

Less than two years after discovery of a mathematical procedure that Bell Labs said could solve a broad range of complex business problems 50 to 100 times faster than current methods, AT&T is filing for patents covering its use. The Karmarkar algorithm, which drew headlines when discovered by researcher Narendra Karmarkar, will be applied first to AT&T's long-distance network.

Thus far, Bell Labs has verified the procedure's capabilities in developing plans for new fiber-optic transmission and satellite capacity linking 20 countries bordering the Pacific Ocean. That jointly owned network will be built during the next 10 years. Planning requires a tremendous number of "what if" scenarios involving 43,000 variables describing transmission capacity, location and construction schedules, all juggled amid political considerations of each connected country.

The Karmarkar algorithm was able to solve the Pacific Basin problem in four minutes, against 80 minutes by the method previously used, says Neil Dinn, head of Bell Labs' international transmission planning department. The speedier solutions will enable international committees to agree on network designs at one meeting instead of many meetings stretched out over months.

AT&T now is using the Karmarkar procedure to plan construction for its domestic network, a problem involving 800,000 variables. In addition, the procedure may be written into software controlling routing of domestic phone calls, boosting the capacity of AT&T's current network.

THE STARTLING DISCOVERY BELL LABS KEPT IN THE SHADOWS

Now its breakthrough mathematical formula could save business millions

It happens all too often in science. An obscure researcher announces a stunning breakthrough and achieves instant fame. But when other scientists try to repeat his results, they fail. Fame quickly turns to notoriety, and eventually the episode is all but forgotten.

That seemed to be the case with Narendra K. Karmarkar, a young scientist at AT&T Bell Laboratories. In late 1984 the 28-year-old researcher astounded not only the scientific community but also the business world. He claimed he had cracked one of the thorniest aspects of computer-aided problem-solving. If so, his feat would have meant an instant windfall for many big companies. It could also have pointed to better software for small companies that use computers to help manage their business.

Karmarkar said he had discovered a quick way to solve problems so hideously complicated that they often defy even the most powerful supercomputers. Such problems bedevil a broad range of business activities, from assessing risk factors in stock portfolios to drawing up production schedules in factories. Just about any company that distributes products through more than a handful of warehouses bumps into such problems when calculating the cheapest routes for getting goods to customers. Even when the problems aren't terribly complex, solving them can chew up so much computer time that the answer is useless before it's found.

HEAD START. To most mathematicians, Karmarkar's precocious feat was hard to swallow. Because such questions are so common, a special branch of mathematics called linear programming (LP) has evolved, and most scientists thought that was as far as they could go. Sure enough, when other researchers independently tried to test Karmarkar's process, their results were disappointing. At scientific conferences skeptics attacked the algorithm's validity as well as Karmarkar's veracity.

But this story may end with a different twist. Other scientists weren't able to duplicate Karmarkar's work, it turns out, because his employer wanted it that way. Vital details about how best to translate the lgorithm, whose mathematical notations run on for about 20 printed pages, into digital computer code were withheld to give Bell Labs a head start at developing commercial products. Following the breakup of American Telephone & Telegraph Co. in January, 1984, Bell Labs was no longer prevented from exploiting its research for profit. While the underlying concept could not be patented or copyrighted because it is pure knowledge, any computer programs that AT&T developed to implement the procedure can be protected.

Now, AT&T may soon be selling the first product based on Karmarkar's work—to the U.S. Air Force. It includes a multiprocessor computer from Alliant Computer Systems Corp. and a software version of Karmarkar's algorithm that has been optimized for high-speed parallel processing. The system would be installed at St. Louis' Scott Air Force Base, headquarters of the Military Airlift Command (MAC). Neither party will comment on the deal's cost or where the negotiations stand, but the Air Force's interest is easy to fathom.

JUGGLING ACT. On a typical day thousands of planes ferry cargo and passengers among air fields scattered around the world. To keep those jets flying, MAC

KARMARKAR: SKEPTICS ATTACKED HIS PRECOCIOUS FEAT

BUSINESS WEEK, September 21, 1987

Simplex Methods versus Interior-point Methods

- *The simplex methods* move from one extreme-point to an **adjacent** one, and they **always stay on the boundary** of the feasible region
- *The Interior-point methods* use **strict interior points**, but **never points on the boundary**
- *The simplex methods* use **algebraic** ideas to move from one points to the next, while *the interior-point methods* use **calculus** (based on derivative)

The Path-following Method

- The path-following method can begin from a point that is neither primal nor dual feasible, and it will proceed from there directly to the optimal solution
- Hence, we start with an arbitrary choice of **strictly positive values** for all the primal and dual variables, i.e., $(x, \omega, y, z) > 0$, and then iteratively update these values as follows:

1) Estimate an appropriate value for μ (i.e., smaller than the "current" value but not too small)
2) Compute **step directions** $(\Delta x, \Delta\omega, \Delta y, \Delta z)$ pointing approximately at the point $(x_\mu, \omega_\mu, y_\mu, z_\mu)$ on the central path
3) Compute a **step length** parameter θ such that the new point:

$$\tilde{x} = x + \theta\Delta x \quad , \quad \tilde{y} = y + \theta\Delta y$$

$$\tilde{\omega} = \omega + \theta\Delta\omega \quad , \quad \tilde{z} = z + \theta\Delta z$$

continues to have strictly positive components

4) Replace (x, ω, y, z) with the new solution $(\tilde{x}, \tilde{\omega}, \tilde{y}, \tilde{z})$

Step Direction

- Our aim is to find step directions, (Δx, $\Delta \omega$, Δy, Δz), such that the new point ($x+\Delta x$, $\omega+\Delta\omega$, $y+\Delta y$, $z+\Delta z$) lies on the **primal–dual central path** (i.e., the path $\{(x_\mu, \omega_\mu, y_\mu, z_\mu) : \mu > 0\}$)
- Recall that, the defining equations for this point on the *central path*:

$$
\begin{aligned}
&Ax + \omega = b \\
&A^T y - z = c \\
&XZe = \mu e \\
&YWe = \mu e
\end{aligned}
$$

- If the new point ($x+\Delta x$, $\omega+\Delta\omega$, $y+\Delta y$, $z+\Delta z$) lies exactly on the central path at μ defined by:

$$
\begin{aligned}
&A(x+\Delta x) + (\omega + \Delta\omega) = b \\
&A^T(y+\Delta y) - (z+\Delta z) = c \\
&(X+\Delta X)(Z+\Delta Z)e = \mu e \\
&(Y+\Delta Y)(W+\Delta W)e = \mu e
\end{aligned}
$$

Step Direction

- (x, ω, y, z) are constants(given), and (Δx, $\Delta\omega$, Δy, Δz) are unknowns, rewriting these equations with the unknowns on the left and the constants on the right:

$$
\begin{aligned}
&A\Delta x + \Delta\omega = b - Ax - \omega =: \rho \\
&A^T\Delta y - \Delta z = c - A^T y + z =: \sigma \\
&Z\Delta x + X\Delta z + \Delta X\Delta Z e = \mu e - XZe \\
&W\Delta y + Y\Delta\omega + \Delta Y\Delta W e = \mu e - YWe
\end{aligned}
$$

,where ρ and σ are vectors represent the primal infeasibility and the dual infeasibility, respectively

Step Direction

- Rearrange with step direction variables on left and drop nonlinear terms on left:

$$
\begin{aligned}
&A\Delta x + \Delta\omega = b - Ax - \omega \\
&A^T\Delta y - \Delta z = c - A^T y + z \\
&Z\Delta x + X\Delta z = \mu e - ZXe \\
&W\Delta y + Y\Delta\omega = \mu e - WYe
\end{aligned}
$$

- This is a linear system of $2m + 2n$ equations in $2m + 2n$ unknowns

The Path-following Method

$$
\begin{array}{l}
\text{initialize}(x,\omega,y,z) > 0 \\
\text{while(not optimal)}\{ \\
\qquad \rho = b - Ax - \omega \\
\qquad \sigma = c - A^T y + z \\
\qquad \gamma = z^T x + y^T \omega \\
\qquad \mu = \delta \dfrac{\gamma}{n+m} \\
\qquad \text{slove}: \\
\qquad\qquad \begin{aligned} A\Delta x + \Delta\omega &= \rho \\ A^T\Delta y - \Delta z &= \sigma \\ Z\Delta x + X\Delta z &= \mu e - XZe \\ W\Delta y + Y\Delta\omega &= \mu e - YWe \end{aligned} \\
\qquad \theta = r\left(\max_{ij}\left\{-\dfrac{\Delta x_j}{x_j}, -\dfrac{\Delta\omega_i}{\omega_i}, -\dfrac{\Delta y_i}{y_i}, -\dfrac{\Delta z_j}{z_j}\right\}\right)^{-1} \wedge \quad 1 \\
\qquad x \leftarrow x + \theta\Delta x, \quad \omega \leftarrow \omega + \theta\Delta\omega \\
\qquad y \leftarrow y + \theta\Delta y, \quad z \leftarrow z + \theta\Delta z \\
\}
\end{array}
$$

Example

- Given the LP:

$$\begin{aligned} \text{maximize} \quad & 2x_1 - 6x_2 \\ \text{subject to} \quad & -x_1 - x_2 - x_3 \le -2 \\ & 2x_1 - x_2 + x_3 \le 1 \\ & x_1, x_2, x_3 \ge 0 \end{aligned}$$

Example

- The dual is :

$$
\begin{aligned}
\text{minimize} \quad & -2y_1 + y_2 \\
\text{subject to} \quad & -y_1 + 2y_2 \geq 2 \\
& -y_1 - y_2 \geq -6 \\
& -y_1 + y_2 \geq 0 \\
& y_1, y_2 \geq 0
\end{aligned}
$$

Example

- Start the algorithm with $(x, w, y, z) = (e, e, e, e)$, and using $\delta = 1/10$, and $r = 9/10$:

$$\rho = b - Ax - \omega$$

$$\begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} - \begin{bmatrix} -1 & -1 & -1 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$$

$$\sigma = c - A^T y + z$$

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -6 \\ 0 \end{bmatrix} - \begin{bmatrix} -1 & 2 \\ -1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$$

$$\gamma = z^T x + y^T \omega = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 5$$

$$\mu = \delta \frac{\gamma}{n+m} = \frac{1}{10} \frac{5}{3+2} = \frac{1}{10}$$

$$\text{initialize}(x,\omega,y,z) > 0$$

$$\text{while(not optimal)}\{$$

$$\boxed{\begin{aligned} \rho &= b - Ax - \omega \\ \sigma &= c - A^T y + z \\ \gamma &= z^T x + y^T \omega \\ \mu &= \delta \frac{\gamma}{n+m} \end{aligned}}$$

slove:

$$\begin{aligned} A\Delta x + \Delta\omega &= \rho \\ A^T \Delta y - \Delta z &= \sigma \\ Z\Delta x + X\Delta z &= \mu e - XZe \\ W\Delta y + Y\Delta\omega &= \mu e - YWe \end{aligned}$$

$$\theta = r\left(\max_{ij}\left\{-\frac{\Delta x_j}{x_j}, -\frac{\Delta\omega_i}{\omega_i}, -\frac{\Delta y_i}{y_i}, -\frac{\Delta z_j}{z_j}\right\}\right)^{-1} \wedge 1$$

$$x \leftarrow x + \theta\Delta x, \quad \omega \leftarrow \omega + \theta\Delta\omega$$

$$y \leftarrow y + \theta\Delta y, \quad z \leftarrow z + \theta\Delta z$$

$$\}$$

Example

$$Compute \quad step \quad direction:$$

$$A\Delta x + \Delta\omega = \rho$$

$$\begin{bmatrix} -1 & -1 & -1 \\ 2 & -1 & 1 \end{bmatrix}\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{bmatrix} + \begin{bmatrix} \Delta\omega_1 \\ \Delta\omega_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$$

$$A^T\Delta y - \Delta z = \sigma$$

$$\begin{bmatrix} -1 & 2 \\ -1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} \Delta y_1 \\ \Delta y_2 \end{bmatrix} - \begin{bmatrix} \Delta z_1 \\ \Delta z_2 \\ \Delta z_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$$

$$Z\Delta x + X\Delta z = \mu e - XZe$$

$$\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{bmatrix} + \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}\begin{bmatrix} \Delta z_1 \\ \Delta z_2 \\ \Delta z_3 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$W\Delta y + Y\Delta\omega = \mu e - YWe$$

$$\begin{bmatrix} 1 & \\ & 1 \end{bmatrix}\begin{bmatrix} \Delta y_1 \\ \Delta y_2 \end{bmatrix} + \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}\begin{bmatrix} \Delta\omega_1 \\ \Delta\omega_2 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}\begin{bmatrix} 1 & \\ & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

initialize$(x,\omega,y,z) > 0$

while(not optimal){

$$\rho = b - Ax - \omega$$

$$\sigma = c - A^T y + z$$

$$\gamma = z^T x + y^T \omega$$

$$\mu = \delta\frac{\gamma}{n+m}$$

slove:

$$\boxed{\begin{aligned} A\Delta x + \Delta\omega &= \rho \\ A^T\Delta y - \Delta z &= \sigma \\ Z\Delta x + X\Delta z &= \mu e - XZe \\ W\Delta y + Y\Delta\omega &= \mu e - YWe \end{aligned}}$$

$$\theta = r\left(\max_{ij}\left\{-\frac{\Delta x_j}{x_j}, -\frac{\Delta\omega_i}{\omega_i}, -\frac{\Delta y_i}{y_i}, -\frac{\Delta z_j}{z_j}\right\}\right)^{-1} \wedge 1$$

$$x \leftarrow x + \theta\Delta x, \quad \omega \leftarrow \omega + \theta\Delta\omega$$

$$y \leftarrow y + \theta\Delta y, \quad z \leftarrow z + \theta\Delta z$$

}

Example

Solve the following system of equations :

$$
\begin{cases}
-\Delta x_1 - \Delta x_2 - \Delta x_3 + \Delta \omega_1 = 0 \\
2\Delta x_1 - \Delta x_2 + \Delta x_3 + \Delta \omega_2 = -2 \\
-\Delta y_1 + 2\Delta y_2 - \Delta z_1 = 2 \\
-\Delta y_1 - \Delta y_2 - \Delta z_2 = -3 \\
-\Delta y_1 + \Delta y_2 - \Delta z_3 = 1 \\
\Delta x_1 + \Delta z_1 = -0.9 \\
\Delta x_2 + \Delta z_2 = -0.9 \\
\Delta x_3 + \Delta z_3 = -0.9 \\
\Delta y_1 + \Delta \omega_1 = -0.9 \\
\Delta y_2 + \Delta \omega_2 = -0.9
\end{cases}
$$

$$
\begin{aligned}
&\text{initialize}(x,\omega,y,z) > 0 \\
&\text{while(not optimal)}\{ \\
&\quad \rho = b - Ax - \omega \\
&\quad \sigma = c - A^T y + z \\
&\quad \gamma = z^T x + y^T \omega \\
&\quad \mu = \delta \frac{\gamma}{n+m} \\
&\quad \text{slove:} \\
&\qquad \boxed{\begin{aligned} A\Delta x + \Delta\omega &= \rho \\ A^T\Delta y - \Delta z &= \sigma \\ Z\Delta x + X\Delta z &= \mu e - XZe \\ W\Delta y + Y\Delta\omega &= \mu e - YWe \end{aligned}} \\
&\quad \theta = r\left(\max_{ij}\left\{-\frac{\Delta x_j}{x_j}, -\frac{\Delta\omega_i}{\omega_i}, -\frac{\Delta y_i}{y_i}, -\frac{\Delta z_j}{z_j}\right\}\right)^{-1} \wedge 1 \\
&\quad x \leftarrow x + \theta\Delta x, \quad \omega \leftarrow \omega + \theta\Delta\omega \\
&\quad y \leftarrow y + \theta\Delta y, \quad z \leftarrow z + \theta\Delta z \\
&\}
\end{aligned}
$$

Example

We got :

$$\begin{cases} \Delta x_1 = -1/2 \\ \Delta x_2 = -14/10 \\ \Delta x_3 = -4/30 \\ \Delta \omega_1 = -61/30 \\ \Delta \omega_2 = -68/30 \\ \Delta z_1 = -4/10 \\ \Delta z_2 = 1/2 \\ \Delta z_3 = -23/30 \\ \Delta y_1 = 34/30 \\ \Delta y_2 = 41/30 \end{cases}$$

Then calculate step length :

$$\theta = r\left(\max_{ij}\left\{-\frac{\Delta x_j}{x_j}, \quad -\frac{\Delta \omega_j}{\omega_j}, \quad -\frac{\Delta y_j}{y_j}, \quad -\frac{\Delta z_j}{z_j}\right\}\right)^{-1} \wedge 1 = \left\{\frac{9}{10}\left(\frac{68}{30}\right)^{-1}\right\} \wedge 1 = \frac{27}{68}$$

initialize$(x,\omega,y,z) > 0$

while(not optimal){

$\rho = b - Ax - \omega$

$\sigma = c - A^T y + z$

$\gamma = z^T x + y^T \omega$

$\mu = \delta \dfrac{\gamma}{n+m}$

slove:

$$\begin{aligned} A\Delta x + \Delta\omega &= \rho \\ A^T\Delta y - \Delta z &= \sigma \\ Z\Delta x + X\Delta z &= \mu e - XZe \\ W\Delta y + Y\Delta\omega &= \mu e - YWe \end{aligned}$$

$$\boxed{\theta = r\left(\max_{ij}\left\{-\frac{\Delta x_j}{x_j}, -\frac{\Delta \omega_i}{\omega_i}, -\frac{\Delta y_i}{y_i}, -\frac{\Delta z_j}{z_j}\right\}\right)^{-1} \wedge 1}$$

$x \leftarrow x + \theta\Delta x, \quad \omega \leftarrow \omega + \theta\Delta\omega$

$y \leftarrow y + \theta\Delta y, \quad z \leftarrow z + \theta\Delta z$

}

Example

$Update \quad (x, \omega, y, z):$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \omega_1 \\ \omega_2 \\ z_1 \\ z_2 \\ z_3 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{27}{68} \begin{bmatrix} -1/2 \\ -14/10 \\ -4/30 \\ -61/30 \\ -68/30 \\ -4/10 \\ 1/2 \\ -23/30 \\ 41/30 \\ 34/30 \end{bmatrix}$$

initialize$(x,\omega,y,z) > 0$

while(not optimal){

$\rho = b - Ax - \omega$

$\sigma = c - A^T y + z$

$\gamma = z^T x + y^T \omega$

$\mu = \delta \dfrac{\gamma}{n+m}$

slove:

$$\begin{aligned} A\Delta x + \Delta\omega &= \rho \\ A^T\Delta y - \Delta z &= \sigma \\ Z\Delta x + X\Delta z &= \mu e - XZe \\ W\Delta y + Y\Delta\omega &= \mu e - YWe \end{aligned}$$

$$\theta = r\left(\max_{ij}\left\{-\frac{\Delta x_j}{x_j}, -\frac{\Delta\omega_i}{\omega_i}, -\frac{\Delta y_i}{y_i}, -\frac{\Delta z_j}{z_j}\right\}\right)^{-1} \wedge \ 1$$

$$\boxed{\begin{aligned} &x \leftarrow x + \theta\Delta x, \quad \omega \leftarrow \omega + \theta\Delta\omega \\ &y \leftarrow y + \theta\Delta y, \quad z \leftarrow z + \theta\Delta z \end{aligned}}$$

}

Convergence Analysis

Questions to Check Whether an Algorithm is Convergent or Not?

- Does the sequence of solutions produced by the path-following method converge?
- If so, is the limit optimal?
- How fast is the convergence?

Measure the Progress of the Algorithm

- Three criteria that must be met to have an optimal solution:
  (1) *Primal feasibility*, (2) *Dual feasibility*, and (3) *Complementarity*
- In order to measure the size of the vector, we may use:
  - The *p-norm* of vector defined as:

$$\|x\|_p = \left(\sum_j |x_j|^p\right)^{\frac{1}{p}}$$

    - *e.g.* The *1-norm*
    - *e.g.* The *sup-norm* is defined by the limit as p tends to infinity: $\|x\|_\infty = \max_j |x_j|.$

Optimal Criteria

- For the primal feasibility criterion, we use the *1-norm* of the primal infeasibility vector:

$$\rho = b - Ax - w.$$

- For the dual feasibility criterion, we use the *1-norm* of the dual infeasibility vector:

$$\sigma = c - A^T y + z.$$

- For complementarity, we use:

$$\gamma = z^T x + y^T w.$$

Progress in One Iteration

- Recall the step size in each iteration is defined as:

$$\theta = r\left(\max_{i,j}\left\{\left|\frac{\Delta x_j}{x_j}\right|, \left|\frac{\Delta w_i}{w_i}\right|, \left|\frac{\Delta y_i}{y_i}\right|, \left|\frac{\Delta z_j}{z_j}\right|\right\}\right)^{-1} \wedge 1$$

- , which is equivalent to :

$$= \frac{r}{\max(\|X^{-1}\Delta x\|_\infty, \ldots, \|Z^{-1}\Delta z\|_\infty)} \wedge 1.$$

Primal and Dual Infeasibilities

- At the next iteration, we have:

$$\tilde{x} = x + \theta \Delta x, \qquad \tilde{y} = y + \theta \Delta y,$$
$$\tilde{w} = w + \theta \Delta w, \qquad \tilde{z} = z + \theta \Delta z.$$

- , and the primal infeasibility is:

$$\begin{aligned}\tilde{\rho} &= b - A\tilde{x} - \tilde{w} \\ &= b - Ax - w - \theta(A\Delta x + \Delta w).\end{aligned}$$

- Since $b - Ax - \omega = \rho$ and $A\,\Delta x + \Delta\omega = \rho$, we obtain:

$$\boxed{\tilde{\rho} = (1 - \theta)\rho.}$$

- Similarly, the dual infeasibility at the next iteration is:

$$\boxed{\begin{aligned}\tilde{\sigma} &= c - A^T\tilde{y} + \tilde{z} \\ &= c - A^T y + z - \theta(A\Delta y - \Delta z) \\ &= (1 - \theta)\sigma.\end{aligned}}$$

Complementarity

- The infeasibility of complementarity at next iteration is:

$$\begin{aligned}\tilde{\gamma} &= \tilde{z}^T\tilde{x} + \tilde{y}^T\tilde{w} \\ &= (z+\theta\Delta z)^T(x+\theta\Delta x) + (y+\theta\Delta y)^T(w+\theta\Delta w) \\ &= z^Tx + y^Tw \\ &\quad + \theta(\boxed{z^T\Delta x + \Delta z^T x} + \boxed{y^T\Delta w + \Delta y^T w}) \\ &\quad + \theta^2(\boxed{\Delta z^T\Delta x + \Delta y^T\Delta w}).\end{aligned}$$

$$\begin{aligned}z^T\Delta x + \Delta z^T x &= e^T(Z\Delta x + X\Delta z) \\ &= e^T(\mu e - ZXe) \\ &= \mu n - z^Tx.\end{aligned}$$

$$\begin{aligned}y^T\Delta w + \Delta y^T w &= e^T(Y\Delta w + W\Delta y) \\ &= e^T(\mu e - YWe) \\ &= \mu m - y^Tw.\end{aligned}$$

$$\begin{aligned}\Delta z^T\Delta x + \Delta y^T\Delta w &= \left(A^T\Delta y - \sigma\right)^T\Delta x + \Delta y^T(\rho - A\Delta x) \\ &= \Delta y^T\rho - \sigma^T\Delta x.\end{aligned}$$

- Thus,

$$\begin{aligned}\tilde{\gamma} &= z^Tx + y^Tw \\ &\quad + \theta\left(\mu(n+m) - (z^Tx + y^Tw)\right) \\ &\quad + \theta^2\left(\Delta y^T\rho - \sigma^T\Delta x\right).\end{aligned}$$

- The γ defined as $z^Tx + y^T\omega = \gamma$ and $\mu\ (n+m) = \delta\gamma$, hence:

$$\boxed{\tilde{\gamma} = (1-(1-\delta)\theta)\gamma + \theta^2\left(\Delta y^T\rho - \sigma^T\Delta x\right).}$$

Apply the Hölder's Inequality

- The Hölder's *inequality* tells us:

$$
\begin{aligned}
|v^T w| &= |\sum_j v_j w_j| \\
&\leq \sum_j |v_j||w_j| \\
&\leq (\max_j |v_j|)(\sum_j |w_j|) \\
&= \|v\|_\infty \|w\|_1 .
\end{aligned}
$$

- Applying the *inequality,* we have:

$$
|\Delta y^T \rho| \leq \|\rho\|_1 \|\Delta y\|_\infty \qquad \text{and} \qquad |\sigma^T \Delta x| \leq \|\sigma\|_1 \|\Delta x\|_\infty .
$$

- Hence,

$$
\tilde{\gamma} \leq (1 - (1 - \delta)\theta)\, \gamma + \theta \left(\|\rho\|_1 \|\theta \Delta y\|_\infty + \|\sigma\|_1 \|\theta \Delta x\|_\infty \right) .
$$

The Bounds of Step Size

- The step size is:

$$\theta = \frac{r}{\max(\|X^{-1}\Delta x\|_\infty, \ldots, \|Z^{-1}\Delta z\|_\infty)} \wedge 1.$$

- Implies:

$$\theta \le \frac{r}{\|X^{-1}\Delta x\|_\infty} \le \frac{x_j}{|\Delta x_j|} \qquad \text{for all } j.$$

- Hence,

$$\|\theta \Delta x\|_\infty \le \|x\|_\infty.$$

- Similarly,

$$\|\theta \Delta y\|_\infty \le \|y\|_\infty.$$

- Finally,

$$\tilde{\gamma} \le (1 - (1 - \delta)\theta)\,\gamma + M\|\rho\|_1 + M\|\sigma\|_1.$$

Stopping Rule

- Let $M < \infty$ and $\varepsilon > 0$,
  - If $\|x\|_{\infty} > M$, then the primal is unbounded
  - If $\|y\|_{\infty} > M$, then the dual is unbounded
  - If $\|\rho\|_{1} < \varepsilon$, $\|\sigma\|_{1} < \varepsilon$ *and* $\gamma < \varepsilon$ then we stop and the current solution is optimal
- Furthermore, let's look at the infeasibility of complementarity:

$$\begin{aligned}\gamma &= z^T x + y^T w \\ &= (\sigma + A^T y - c)^T x + y^T (b - Ax - \rho) \\ &= b^T y - c^T x + \sigma^T x - \rho^T y.\end{aligned}$$

- Applying the Hölder's inequality, the duality gap is:

$$\begin{aligned}|b^T y - c^T x| &\le \gamma + |\sigma^T x| + |y^T \rho| \\ &\le \gamma + \|\sigma\|_1 \|x\|_\infty + \|\rho\|_1 \|y\|_\infty.\end{aligned}$$

- The above inequality implies that the dual gap is very small as $\|\rho\|_{1}$, $\|\sigma\|_{1}$ and γ are small

The Convergence of the Path-following Algorithm

Theorem: Suppose there exists t > 0, $M < \infty$, and an integer K such that $k \le K$, $\theta^{(k)} \ge t$, $\| x^{(k)} \|_\infty > M$, and $\| y^{(k)} \|_\infty > M$. Then there exists a constant $\bar{M} < \infty$ such that:

$$
\begin{aligned}
\|\rho^{(k)}\|_1 &\le (1-t)^k \|\rho^{(0)}\|_1, \\
\|\sigma^{(k)}\|_1 &\le (1-t)^k \|\sigma^{(0)}\|_1, \\
\gamma^{(k)} &\le (1-\tilde{t})^k \bar{M},
\end{aligned}
$$

for all $k \le K$, where $\tilde{t} = t(1-\delta)$.

Homework

- Exercise 18.1 (part a)
- Exercise 18.6 (*Higher-order methods*)