LECTURE NOTE VI

Chapter 10

Quadratic Approximation Methods for Constrained Problems

- Linear approximation can be applied to a constrained NLP and the LP subproblem suggested a corner point of the constraints. However, the improved feasible point may be at the corner or inside the feasible region.
- Thus, a line search will be performed to find the improved point in the direction to the suggested points by the LP subproblem.
- A better way to approximate the nonlinear objective function and constraints is higher-order approximation such as quadratic approximation.

10.1 Direct Quadratic Approximation

▶ Quadratic approximation of a function

$$q(x;x^0)=f(x^0)+\nabla f(x^0)^T(x-x^0)+\frac{1}{2}(x-x^0)^T\nabla^2 f(x^0)(x-x^0)$$

▶ Direct Successive Quadratic programming solution

Given x^0, an initial solution estimate, and a suitable method for solving QP subproblems.

Step 1: Formulate the QP problem:

$$\min \nabla f(x^{(k)})^T d+\frac{1}{2}d^T\nabla^2 f(x^{(k)})d$$

$$\text{subject to } h_i(x^{(k)})+\nabla h_i(x^{(k)})^T d=0 \quad (i=1,\cdots,I)$$

$$g_j(x^{(k)})+\nabla g_j(x^{(k)})^T d=0 \quad (j=1,\cdots,J)$$

Step 2: Solve the QP problem and set $x^{(k+1)}= x^{(k)}+d$.

Step 3: Check the convergence. If not converged, repeat step 1.

Example: Favorable case

$$\min\ f(x)=6x_1x_2^{-1}+x_2x_1^{-2}$$

$$\text{subject to } h(x)=x_1x_2-2=0$$

$$g(x)=x_1+x_2-1\ge 0$$

From an initial guess $x^0=(2\ 1)^T$ and $f(x^0)=12.25$, using the direct successive QP strategy.

$$\nabla f(x)=(6x_2^{-1}-2x_2x_1^{-3}\quad -6x_1x_2^{-2}+x_1^{-2})^T$$

$$\nabla^2 f(x)=\begin{pmatrix} 6x_2x_1^{-4} & -6x_2^{-2}-2x_1^{-3} \\ -6x_2^{-2}-2x_1^{-3} & 12x_1x_2^{-3}\end{pmatrix}$$

$$\nabla h(x)=(x_2\quad x_1)^T$$

Thus, the first QP subproblem will be

$$\min\ f(x)=(23/4\quad -47/4)d+\frac{1}{2}d^T\begin{pmatrix} 3/8 & -25/4 \\ -25/4 & 24\end{pmatrix}d$$

$$\text{subject to } (1\quad 2)d=0$$

$$(1\quad 1)d+2\ge 0$$

The solution is $d^0=(-0.92079\quad 0.4604)^T$ and $x^{(1)}= x^0+d^0$. At this point

$f(x^{(1)})=5.68779$ (improved), $h(x^{(1)})= -0.42393$ (not feasible), $g(x^{(1)})\ge 0$

Then, the second QP subproblem will be

$$\min\ f(x) = (1.78475 \quad -2.1775)d + \frac{1}{2}d^T \begin{pmatrix} 6.4595 & -4.4044 \\ -4.4044 & 4.1579 \end{pmatrix} d$$

$$\text{subject to } (1.4604 \quad 1.07921)d - 0.42393 = 0$$

$$(1 \quad 1)d + 1.5396 \geq 0$$

The solution is $d^{(1)}$=(–0.03043 0.434)^T and $x^{(2)} = x^{(1)} + d^{(1)}$. At this point

$f(x^{(2)})$=5.044 (improved), $h(x^{(2)})$= –0.0132 (reduced violation), $g(x^{(2)})\geq 0$

The next two iterations produce

$x^{(3)}$=(1.00108 1.99313)^T, $f(x^{(3)})$=5.00457, $h(x^{(3)})$= –4.7x10^{-3}, $g(x^{(3)})\geq 0$

$x^{(4)}$=(1.00014 1.99971)^T, $f(x^{(4)})$=5.00003, $h(x^{(4)})$= –6.2x10^{-6}, $g(x^{(4)})\geq 0$

In 4 iterations, a very accurate solution has been obtained. (x^*=(1 2))

Example: Unfavorable case

$$\min\ f(x) = x_1 x_2$$

$$\text{subject to } h(x) = 6x_1x_2^{-1} + x_2x_1^{-2} - 5 = 0$$

$$g(x) = x_1 + x_2 - 1 \geq 0$$

From an initial guess x^0=(2 1)^T and $f(x^0)$=12.25, using the direct successive QP strategy.

$$\nabla f(x) = (x_2 \quad x_1)^T$$

$$\nabla^2 f(x) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\nabla h(x) = (6x_2^{-1} - 2x_2x_1^{-3} \quad -6x_1x_2^{-2} + x_1^{-2})^T$$

Thus, the first QP subproblem will be

$$\min\ f(x) = (1 \quad 2)d + \frac{1}{2}d^T \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} d$$

$$\text{subject to } (23/4 \quad -47/4)d + 29/4 = 0$$

$$(1 \quad 1)d + 2 \geq 0$$

The solution is d^0=(–1.7571 –0.24286)^T and $x^{(1)} = x^0 + d^0$. At this point

$f(x^{(1)})$=0.18388, $h(x^{(1)})$= 9.7619, $g(x^{(1)})$=0

Then, the second QP subproblem will be

$$\min\ f(x) = (0.75714 \quad 0.24286)d + \frac{1}{2}d^T \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} d$$

$$\text{subject to } (-97.795 \quad 14.413)d + 9.7619 = 0$$

$$(1 \quad 1)d \geq 0$$

The resulting new points for the next 4 iterations are:

$x^{(2)}$=(0.32986 0.67015)^T, $f(x^{(2)})$=0.2211, $h(x^{(2)})$= 4.1123, $g(x^{(2)})$=0

$x^{(3)}$=(0.45383 0.54618)^T, $f(x^{(3)})$=0.2479, $h(x^{(3)})$= 2.6374, $g(x^{(3)})$=0

$x^{(4)}$=(–0.28459 1.28459)^T, $f(x^{(4)})$= –0.3656, $h(x^{(4)})$= 9.5316, $g(x^{(4)})$=0

$x^{(5)}$=(–0.19183 1.19183)^T, $f(x^{(5)})$= –0.2286, $h(x^{(5)})$= 26.422, $g(x^{(5)})$=0

The iterations show oscillatory path. (x^*=(1 2)^T) and the feasibility is not improving.

➔ The nonlinear constraints have to be handled appropriately.

10.2 Quadratic Approximation of the Lagrangian Function

▶ Equality Constrained Case

min $f(x)$

subject to $h(x)$ =0

- Necessary condition for optimality of the Lagrangian function

$$\nabla_x L(x^*, v^*) = \nabla f^* - v^{*T} \nabla h^* = 0 \quad \text{and} \quad h(x^*) = 0$$

- Sufficient condition for optimality

$$\nabla_x^2 L(x^*, v^*) = \nabla^2 f^* - v^{*T} \nabla^2 h^* \quad \text{satisfy} \quad d^T \nabla_x^2 L d > 0 \quad \forall d \ni \nabla h^{*T} d = 0$$

▶ QP subproblem

$$\min \nabla f(x^{(k)})^T d + \frac{1}{2} d^T \nabla_x^2 L(x^{(k)}, v^{(k)}) d$$

$$\text{subject to } h_i(x^{(k)}) + \nabla h_i(x^{(k)})^T d = 0 \quad (i = 1, \cdots, I)$$

- If d^*=0 is the solution of this subproblem, $x^{(k)}$ must satisfy the necessary conditions for a local minimum of the original problem. Then the Lagrange multiplier can be obtained from

$$\nabla f(x^{(k)}) - v^{*T} \nabla h(x^{(k)}) = 0$$

- Suppose this subproblem satisfies the second-order sufficient conditions for optimality at d=0 with v^*. Then

$$y^T \nabla_d^2 \left[\left\{ \nabla f(x^{(k)}) d + 0.5 d^T \nabla_x^2 L(x^{(k)}, v^*) d \right\} - v^{*T} \left\{ h(x^{(k)}) + \nabla h(x^{(k)})^T d \right\} \right] y$$

$$= y^T \nabla_x^2 L(x^{(k)}, v^*) y > 0 \quad \forall y \ni \nabla h(x^{(k)})^T d = 0$$

and ($x^{(k)}$, v^*) satisfies the sufficient conditions for a local minimum of the original problem.

- If no further corrections can be found, that is d=0, then the local minimum of the original problem will have been obtained.
- The Lagrange multiplier of the subproblem can be used conveniently as estimates of the multipliers used to formulate the next subproblem.
- For points sufficiently close to the solution of the original problem the quadratic objective function is likely to be positive definite and thus the solution of the QP subproblem will be well behaved.

▶ General case

$$\min \nabla f(x^{(k)})^T d + \frac{1}{2} d^T \nabla_x^2 L(x^{(k)}, v^{(k)}, u^{(k)}) d$$

$$\text{subject to } h_i(x^{(k)}) + \nabla h_i(x^{(k)})^T d = 0 \quad (i = 1, \cdots, I)$$

$$g_j(x^{(k)}) + \nabla g_j(x^{(k)})^T d \geq 0 \quad (j = 1, \cdots, J)$$

Example: Favorable case

$$\min \ f(x) = 6 x_1 x_2^{-1} + x_2 x_1^{-2}$$

$$\text{subject to } h(x) = x_1 x_2 - 2 = 0$$

$$g(x) = x_1 + x_2 - 1 \geq 0$$

From an initial guess x^0=(2,1), u^0=0 and v^0=0, the first QP subproblem will be

$$\min\ f(x) = (23/4 \quad -47/4)d + \frac{1}{2}d^T \begin{pmatrix} 3/8 & -25/4 \\ -25/4 & 24 \end{pmatrix} d$$

$$\text{subject to } (1 \quad 2)d = 0$$

$$(1 \quad 1)d + 2 \geq 0$$

The solution is d^0=(–0.92079 0.4604)^T. Since $g(d^0)$>0, u^0 must be zero. Also, from the necessary condition, $\nabla f(x^{(k)}) - v^{*T}\nabla h(x^{(k)}) = 0$,

$$(23/4 \quad -47/4) + d^T \begin{pmatrix} 3/8 & -25/4 \\ -25/4 & 24 \end{pmatrix} = v^{(1)}(1 \quad 2)$$

Thus, $v^{(1)}$=2.52723. The new estimate of the problem solution will be $x^{(1)}= x^0+d^0$.

$x^{(1)}$=(1.07921 1.4604)^T, $f(x^{(1)})$=5.68779, $h(x^{(1)})$= –0.42393

$$\nabla f(x^{(1)}) = (1.78475 \quad -2.21775)^T \qquad \nabla h(x^{(1)}) = (1.4604 \quad 1.07921)$$

$$\nabla^2 f(x^{(1)}) = \begin{pmatrix} 6.45924 & -4.40442 \\ -4.40442 & 4.15790 \end{pmatrix} \qquad \nabla^2 h(x^{(1)}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\nabla^2 L = \nabla^2 f - v\nabla^2 h = \nabla^2 f - 2.52723\nabla^2 h = \begin{pmatrix} 6.45924 & -6.93165 \\ -6.93165 & 4.15790 \end{pmatrix}$$

Then, the second subproblem will be

$$\min\ f(x) = (1.78475 \quad -2.21775)d + \frac{1}{2}d^T \begin{pmatrix} 6.45924 & -6.93165 \\ -6.93165 & 4.15790 \end{pmatrix} d$$

$$\text{subject to } (1.4604 \quad 1.07921)d - 0.42393 = 0$$

$$(1 \quad 1)d + 1.539604 \geq 0$$

The solution is $d^{(1)}$=(0.00614 0.3845)^T. Again, since $g(d^{(1)})$>0, $u^{(1)}$ must be zero. Also,

$$(-0.84081 \quad -0.62135) = v^{(2)}(1.4604 \quad 1.07921)$$

Thus, $v^{(2)}$= –0.57574. The new estimate of the problem solution will be

$x^{(2)}$=(1.08535 1.8449)^T, $f(x^{(2)})$=5.09594, $h(x^{(2)})$= 2.36x10^{-3}

The resulting new points for the next 2 iterations are:

$v^{(3)}$= –0.44046, $x^{(3)}$=(0.99266 2.00463)^T, $f(x^{(3)})$=4.99056, $h(x^{(3)})$= –1.008x10^{-2}

$v^{(4)}$= –0.49997, $x^{(4)}$=(0.99990 2.00017)^T, $f(x^{(4)})$=45.00002, $h(x^{(4)})$= –3.23x10^{-5}

The iterations show oscillatory path. (x^*=(1 2)) and the feasibility is not improving.

Example: Unfavorable case

$$\min\ f(x) = x_1 x_2$$

$$\text{subject to } h(x) = 6x_1x_2^{-1} + x_2x_1^{-2} - 5 = 0$$

$$g(x) = x_1 + x_2 - 1 \geq 0$$

From an initial guess x^0=(2 2.789) , u^0=0 and v^0=0, the first QP subproblem will be

$$\min\ f(x) = (2.789 \quad 2)d + \frac{1}{2}d^T \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} d$$

$$\text{subject to } (1.454 \quad -1.2927)d - 1.3\times10^{-4} = 0$$

$$(1 \quad 1)d + 3.789 \geq 0$$

The solution is d^0=(–1.78316 –2.00583)^T. Since the inequality is tight, the multiplier must be calculated.

$$\begin{pmatrix} 0.78317 \\ 0.21683 \end{pmatrix} = v \begin{pmatrix} -145.98 \\ 19.148 \end{pmatrix} + u \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Thus, $v^{(1)}= -0.00343$, $u^{(1)}= 0.28251$, $x^{(1)}= (0.21683\ 0.78317)^T$ and

$f(x^{(1)})=0.1698$, $h(x^{(1)})= 13.318$ (large violation), $g(x^{(1)})=0$

Then, the second QP subproblem will be

$$\min\ f(x) = (0.78317 \quad 0.21683)d$$

$$+\frac{1}{2}d^T \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - v^{(1)} \begin{pmatrix} 2125.68 & -205.95 \\ -205.95 & 5.4168 \end{pmatrix} - u^{(1)} \times 0 \right] d$$

$$\text{subject to } (-145.98 \quad 19.148)d + 13.318 = 0$$

$$(1 \quad 1)d \ge 0$$

Thus, $d^{(1)}= (0.10434\ -0.10434)^T$, $v^{(2)}= -0.02497$, and $u^{(2)}= 0.38822$.

The resulting new points for the next few iterations are:

$x^{(2)}=(0.40183\ 0.59817)^T$, $f(x^{(2)})=0.24036$, $h(x^{(2)})= 2.7352$, $g(x^{(2)})=0$

$x^{(3)}=(0.74969\ 0.25031)^T$, $f(x^{(3)})=0.18766$, $h(x^{(3)})= 13.416$, $g(x^{(3)})=0$

The violation is increased considerably. The convergence is unsatisfactory.

➔ Thus, it suggests that a line search from the previous solution estimate in the direction obtained from the current QP subproblem solution. The line search must be carried using some type of penalty function.

$$P(x,R) = f(x) + R\left\{ \sum_{i=1}^{I} h_i(x)^2 + \sum_{j=1}^{J} \min(0, g_i(x))^2 \right\}$$

From the above example, starting with $x^{(2)}$, $d^{(2)}=(0.34786\ 0.347860)^T$, let

$$P(x,R) = f(x) + 10\left\{ h(x)^2 + \min(0, g(x)^2) \right\}$$

and the minimization along the line

$$x = x^{(2)} + \alpha d^{(2)} = \begin{pmatrix} 0.40183 \\ 0.59817 \end{pmatrix} + \alpha \begin{pmatrix} 0.34786 \\ -0.34786 \end{pmatrix}$$

If $\alpha=0$, $P=75.05$ and if $\alpha=1$, $P=1800$. Using any convenient line search method, the approximate minimum value of $P=68.11$ can be obtained at $\alpha=0.1$. Then the resulting point will be

$x^{(3)}=(0.43662\ 0.56338)^T$, $f(x^{(3)})=0.24682$, $h(x^{(3)})= 2.6053$

However, $d= x^{(3)} - x^{(2)}$ is no longer the optimum solution of the previous subproblem and the only available updated multipliers are associated with $d^{(2)}$, namely $v^{(3)}= -0.005382$ and $u^{(3)}=0.37291$. In this way, the optimal solutions to subproblems will approach to the optimum.

10.3 Variable Metric for Constrained Optimization

- The calculation of Hessian can be approximated using the variable metric methods to reduce the computational burden.

▶ Constrined Variable Metric Method

Given initial estimate x^0, v^0, u^0, and a symmetric positive definite matrix H^0.

Step 1: Solve the problem

$$\min \nabla f(x^{(k)})^T d + \frac{1}{2} d^T H^{(k)} d$$

$$\text{subject to } h_i(x^{(k)}) + \nabla h_i(x^{(k)})^T d = 0 \quad (i = 1, \cdots, I)$$

$$g_j(x^{(k)}) + \nabla g_j(x^{(k)})^T d = 0 \quad (j = 1, \cdots, J)$$

Step 2: Select the step size α along $d^{(k)}$, and set $x^{(k+1)}= x^{(k)}+\alpha d^{(k)}$ to minimize a suitable penalty function.

Step 3: Check the convergence

Step 4: Update $H^{(k)}$, using the gradient difference

$$\nabla_x L(x^{(k+1)}, v^{(k+1)}, u^{(k+1)}) - \nabla_x L(x^{(k)}, v^{(k+1)}, u^{(k+1)})$$

in such a way that $H^{(k+1)}$ remains positive definite. (variable metric method)

Remarks:

1. Penalty functions

   - Han

$$P(x,R) = f(x) + R\left\{\sum_{i=1}^{I}|h_i(x)| + \sum_{j=1}^{J}\min(0, g_i(x))\right\}$$

   - Powell

$$P(x,R) = f(x) + \sum_{i=1}^{I}\mu_i\,|h_i(x)| + \sum_{j=1}^{J}\sigma_j \min(0, g_i(x))$$

where $\mu_i = |v_i|$ and $\sigma_j = |u_j|$ for k=1 and otherwise,

$$\mu_i^{(k)} = \max(|v_i^{(k)}|, 0.5(\mu_i^{(k-1)} + |v_i^{(k)}|))$$

$$\sigma_j^{(k)} = \max(|u_j^{(k)}|, 0.5(\sigma_j^{(k-1)} + |u_j^{(k)}|))$$

2. Hessian update formulae

   - DFP

   - BFS

$$H^{(k+1)} = H^{(k)} - \frac{H^{(k)} z z^T H^{(k)}}{z^T H^{(k)} z} + \frac{w w^T}{z^T w}$$

where $z=x^{(k+1)}-x^{(k)}$, $w=\theta y+(1-\theta)H^{(k)}z$,

$$y = \nabla_x L(x^{(k+1)}, v^{(k+1)}, u^{(k+1)}) - \nabla_x L(x^{(k)}, v^{(k+1)}, u^{(k+1)}) \text{ and}$$

$$\theta = \begin{cases} 1 & \text{if } z^T y \ge 0.2 z^T H^{(k)} z \\ \dfrac{0.8 z^T H^{(k)} z}{z^T H^{(k)} z - z^T y} & \text{otherwise} \end{cases}$$

Example:

$$\min f(x) = 6x_1 x_2^{-1} + x_2 x_1^{-2}$$

$$\text{subject to } h(x) = x_1 x_2 - 2 = 0$$

$$g(x) = x_1 + x_2 - 1 \ge 0$$

From an initial guess x^0=(2,1), and initial metric H^0=I,

$$\nabla f = (23/4 \quad -47/4)^T,\ \nabla h = (1 \quad 2)^T,\ \nabla g = (1 \quad 1)^T$$

Therefore, the first QP subproblem will be

$$\min f(x) = (23/4 \quad -47/4)d + \frac{1}{2}d^T d$$

$$\text{subject to } (1 \quad 2)d = 0$$

$$(1 \quad 1)d + 2 \ge 0$$

The solution is $d^0=(-4 \quad 2)^T$. The multipliers at this point are

$$\begin{pmatrix} 23/4 \\ -47/4 \end{pmatrix} + \begin{pmatrix} -4 \\ 2 \end{pmatrix} = v\begin{pmatrix} 1 \\ 2 \end{pmatrix} + u\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Thus, $v^{(1)}=-46/4$ and $u^{(1)}=53/4$. For the first iteration, let the penalty parameters be

$$\mu^{(1)} = \left|-46/4\right| \quad \text{and} \quad \sigma^{(1)} = \left|53/4\right|$$

The penalty function takes the form (Powell's approach)

$$P = 6x_1x_2^{-1} + x_2x_1^{-2} + (46/4)\left|x_1x_2 - 2\right| - (53/4)\min(0, x_1 + x_2 - 1)$$

The line search in the direction of $x^{(1)}= x^0+\alpha d^0$. Using a bracketing method, P=9.1702 at α=0.1348. The new point is $x^{(1)}=(1.46051\ 1.26974)^T$.

Using BFS update rule,

$z= x^{(1)}- x^0=(-0.53949 \quad 0.26974)^T$

$$\nabla_x L(x^0, v^{(1)}, u^{(1)}) = \begin{pmatrix} 23/4 \\ -47/4 \end{pmatrix} - (-46/4)\begin{pmatrix} 1 \\ 2 \end{pmatrix} - (53/4)\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ -2 \end{pmatrix}$$

$$\nabla_x L(x^{(1)}, v^{(1)}, u^{(1)}) = \begin{pmatrix} 3.91022 \\ -4.9665 \end{pmatrix} - (-46/4)\begin{pmatrix} 1.26974 \\ 1.46051 \end{pmatrix} - (53/4)\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 5.26228 \\ 4.81563 \end{pmatrix}$$

$$y = \nabla_x L(x^{(1)}) - \nabla_x L(x^0) = (1.26228 \quad 6.81563)^T$$

Since z^Ty=1.15749>0.2z^Tz=0.0728, θ=1 and w=y. Thus,

$$H^{(1)} = I - zz^T/\|z\|^2 + yy^T/z^Ty = \begin{pmatrix} 1.57656 & 7.83267 \\ 7.83267 & 40.9324 \end{pmatrix} \quad \text{(p.d.)}$$

Then, the second subproblem will be

$$\min f(x) = (3.91022 \quad -4.9665)d + \frac{1}{2}d^T \begin{pmatrix} 1.57656 & 7.83267 \\ 7.83267 & 40.9324 \end{pmatrix} d$$

$$\text{subject to } (1.26974 \quad 1.46051)d - 0.14552 = 0$$

$$(1 \quad 1)d + 1.73026 \geq 0$$

The solution is $d^{(1)}=(-0.28911 \quad 0.35098)^T$. Since $g(d^{(1)})>0$, $u^{(2)}$ must be zero and $v^{(2)}$=4.8857. The penalty function multipliers are updated.

$$\mu^{(2)} = \max(\left|4.8857\right|, 46/4 + 4.8857/2) = 8.19284$$

$$\sigma^{(2)} = \max(\left|0\right|, (53/4 + 0)/2) = 6.625$$

The penalty function now becomes

$$P(x(\alpha)) = f(x) + 8.19284\left|x_1x_2 - 2\right| - 6.625\min(0, x_1 + x_2 - 1)$$

$$\text{where } x(\alpha) = x^{(1)} + \alpha d^{(1)} \quad 0 \leq \alpha \leq 1$$

The minimum occurs at α=1, $P(x(1))$=6.34906. The new point is

$x^{(2)}=(1.17141\ 1.62073)^T$, $f(x^{(2)})$=5.5177, $h(x^{(2)})=-0.10147$

The iterations continue with an update of $H^{(1)}$. The results are

$v^{(3)}=-0.13036$, $x^{(3)}=(1.14602\ 1.74619)^T$, $f(x^{(3)})$=5.2674, $h(x^{(3)})=-0.13036$

$v^{(4)}=-0.17609$, $x^{(4)}=(1.04158\ 1.90479)^T$, $f(x^{(4)})$=5.0367, $h(x^{(4)})=-0.17090$

$v^{(5)}=-0.45151$, $x^{(5)}=(0.99886\ 1.99828)^T$, $f(x^{(5)})$=5.0020, $h(x^{(4)})=-0.00399$

$v^{(6)}=-0.50128$, $x^{(4)}=(1.00007\ 1.99986)^T$, $f(x^{(4)})$=5.0000, $h(x^{(4)})=-0.1.9\text{x}10^{-6}$

The quasi-Newton iteration that uses only first derivatives is quite satisfactory.

RECOMMENDATIONS FOR CONSTRAINED OPTIMIZATION

1. Best current algorithms
   - GRG 2/CONOPT
   - MINOS
   - SQP
2. GRG 2 (or CONOPT) is generally slower, but is robust. Use with highly nonlinear functions.
   - Solver in Excel
3. For small problems ($n \leq 100$) with nonlinear constraints, use SQP.
4. For large problems ($n \geq 100$) with mostly linear constraints, use MINOS.
   - Difficulty with many nonlinearities

Small, Nonlinear Problems - SQP (generic) solves QP's, not linearly constrained LP's, fewer function calls.

Large, Mostly Linear Problems - MINOS performs sparse constraint decomposition. Works efficiently in reduced space if function calls are cheap.

RULES FOR FORMULATING NONLINEAR PROGRAMS

1) Avoid overflows and undefined terms, (do not divide, take logs, etc.)

   $x + y - \ln(z) = 0 \quad \rightarrow \quad x + y - u = 0$ and $\exp(u) - z = 0$

2) If constraints must always be enforced, make sure they are linear or bounds.

   $v(xy - z^2)^{1/2} = 3 \quad \rightarrow \quad vu = 3,\ u^2 - (xy - z^2) = 0$, and $u \geq 0$

3) Exploit linear constraints as much as possible.

   Mass balance: $x_iL + y_iV = z_iF \quad \rightarrow \quad l_i + v_i = f_i$ and $L - \text{sum}(l_i) = 0$, $V - \text{sum}(v_i) = 0, \ldots$

4) Use bounds and constraints to enforce characteristic solutions.

   If necessary, add constraints such as $a \leq x \leq b$ and $g(x) \leq 0$ to isolate correct root of $h(x) = 0$.

5) Exploit global properties when possibility exists.

   If the problem is convex (no nonlinear equations), then use LP or QP.

   If the problem is a Geometric Program, the logarithmic transformation converts the problem to a convex problem.

6) Exploit problem structure when possible.

   min $(Tx - 3Ty)$

s.t. $xT + y - T^2y = 5$	$\rightarrow$	If T is fixed, it can be solved by LP.
$4x - 5Ty + Tx = 7$		Put T in outer optimization loop.
$0 \leq T \leq 1$		

AVAILABLE SOFTWARE FOR CONSTRAINED OPTIMIZATION

NaG Routines

Unconstrained Optimization

E04CCF - Unconstrained minimum, simplex algorithm, function of several variables using function values only (comprehensive)

E04DGF - Unconstrained minimum, preconditioned conjugate gradient algorithm, function of several variables using first derivatives (comprehensive)

E04FCF - Unconstrained minimum of a sum of squares, combined Gauss-Newton and modified Newton algorithm using function values only (comprehensive)

E04FYF - Unconstrained minimum of a sum of squares, combined Gauss-Newton and modified Newton algorithm using function values only (easy-to-use)

E04GBF - Unconstrained minimum of a sum of squares, combined Gauss-Newton and quasi-Newton algorithm using first derivatives (comprehensive)

E04GDF - Unconstrained minimum of a sum of squares, combined Gauss-Newton and modified Newton algorithm using first derivatives (comprehensive)

E04GYF - Unconstrained minimum of a sum of squares, combined Gauss-Newton and quasi-Newton algorithm, using first derivatives (easy-to-use)

E04GZF - Unconstrained minimum of a sum of squares, combined Gauss-Newton and modified Newton algorithm using first derivatives (easy-to-use)

E04HEF - Unconstrained minimum of a sum of squares, combined Gauss-Newton and modified Newton algorithm, using second derivatives (comprehensive)

E04HYF - Unconstrained minimum of a sum of squares, combined Gauss-Newton and modified Newton algorithm, using second derivatives (easy-to-use)

E04JYF - Minimum, function of several variables, quasi-Newton algorithm, simple bounds, using function values only (easy-to-use)

E04KDF - Minimum, function of several variables, modified Newton algorithm, simple bounds, using first derivatives (comprehensive)

E04KYF - Minimum, function of several variables, quasi-Newton algorithm, simple bounds, using first derivatives (easy-to-use)

E04KZF - Minimum, function of several variables, modified Newton algorithm, simple bounds, using first derivatives (easy-to-use)

E04LBF - Minimum, function of several variables, modified Newton algorithm, simple bounds, using first and second derivatives (comprehensive)

E04LYF - Minimum, function of several variables, modified Newton algorithm, simple bounds, using first and second derivatives (easy-to-use)

Specialized Constrained Algorithms

E04MFF - LP problem (dense)

E04NCF - Convex QP problem or linearly-constrained linear least-squares problem (dense)

SQP Routines

E04UCF - Minimum, function of several variables, sequential QP method, nonlinear constraints, using function values and optionally first derivatives (forward communication, comprehensive)

E04UFF - Minimum, function of several variables, sequential QP method, nonlinear constraints, using function values and optionally first derivatives (reverse communication, comprehensive)

E04UNF - Minimum of a sum of squares, nonlinear constraints, sequential QP method, using function values and optionally first derivatives (comprehensive)

GAMS Programs

CONOPT - Generalized Reduced Gradient method with restoration

MINOS - Generalized Reduced Gradient method with restoration

MS Excel

Solver uses Generalized Reduced Gradient method with restoration