SQP
Sequential Quadratic Programming

This Hypercard stack was prepared by:
Dennis L. Bricker,
Dept. of Industrial Engineering,
University of Iowa,
Iowa City, Iowa 52242
e-mail: dbricker@icaen.uiowa.edu

author

Consider the nonlinear programming problem:

$$\begin{array}{ll} \text{Minimize} & f(x) \\ \text{subject to} & h_k(x) = 0,\ k=1,\cdots K \\ & g_j(x) \le 0,\ j=1,\cdots J \end{array}$$

Given a solution estimate $\bar{x}$ and step d,

$$\begin{cases} f(\bar{x}+d) = f(\bar{x}) + \left[\nabla f(\bar{x})\right]^T d + \tfrac{1}{2}\, d^T \nabla^2 f(\bar{x}) d + \cdots \\ h_k(\bar{x}+d) = h_k(\bar{x}) + \left[\nabla h_k(\bar{x})\right]^T d + \tfrac{1}{2}\, d^T \nabla^2 h_k(\bar{x}) d + \cdots \\ g_j(\bar{x}+d) = g_j(\bar{x}) + \left[\nabla g_j(\bar{x})\right]^T d + \tfrac{1}{2}\, d^T \nabla^2 g_j(\bar{x}) d + \cdots \end{cases}$$

Form the linearly-constrained/quadratic minimization problem:

$$\begin{array}{l} \underset{d}{\text{Minimize}}\ f(\bar{x}) + \left[\nabla f(\bar{x})\right]^T d + \tfrac{1}{2}\, d^T\, \nabla^2 f(\bar{x}) d \\ \text{subject to} \\ \qquad h_k(\bar{x}) + \left[\nabla h_k(\bar{x})\right]^T d = 0,\ k=1,\cdots K \\ \qquad g_j(\bar{x}) + \left[\nabla g_j(\bar{x})\right]^T d \le 0,\ j=1,\cdots J \end{array}$$

EXAMPLE

$$\text{Minimize } f(x) = 6\,\frac{x_1}{x_2} + \frac{x_2}{x_1^2}$$

subject to

$$h(x) = x_1 x_2 - 2 = 0$$
$$g(x) = 1 - x_1 - x_2 \le 0$$

Note that this is a nonconvex problem....
h(x) is nonlinear and f(x) is nonconvex!

$$\nabla f(x) = \begin{bmatrix} \frac{6}{x_2} - 2\,\frac{x_2}{x_1^3} \\ -6\frac{x_1}{x_2^2} + \frac{1}{x_1^2} \end{bmatrix}, \quad \nabla^2 f(x) = \begin{bmatrix} 6\frac{x_2}{x_1^4} & \frac{-6}{x_2^2} - \frac{2}{x_1^3} \\ \frac{-6}{x_2^2} - \frac{2}{x_1^3} & \frac{12 x_1}{x_2^3} \end{bmatrix}$$

$$\nabla h(x) = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}, \qquad \nabla^2 h(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\nabla g(x) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \qquad \nabla^2 g(x) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Let the starting point be $X^o = (2,1)$

$$f(X^o) = 12.25, \quad \nabla f(X^o) = \begin{bmatrix} 23/4 \\ -47/4 \end{bmatrix}, \quad \nabla^2 f(X^o) = \begin{bmatrix} 3/8 & -25/4 \\ -25/4 & 24 \end{bmatrix}$$

$$h(X^o) = 0, \qquad \nabla h(X^o) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \nabla^2 h(X^o) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$g(X^o) = -2 < 0, \qquad \nabla g(X^o) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad \nabla^2 g(X^o) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

(slack)

At $X^o = (2,1)$, the approximating QP is

$$\begin{array}{ll} \text{Minimize} & \left[\nabla f(X^o)\right]^T d + \tfrac{1}{2}\, d^T \nabla^2 f(X^o) d \\ \text{s.t.} & \left[\nabla h(X^o)\right]^T d = -h(X^o) \\ & \left[\nabla g(X^o)\right]^T d \le -g(X^o) \end{array}$$

$$\text{Minimize} \begin{bmatrix} 23/4 & -47/4 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + \tfrac{1}{2} \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 3/8 & -25/4 \\ -25/4 & 24 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

subject to

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} -1 & -1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \le 2$$

This QP problem has the optimal solution:

$$d_1 = -0.92079$$
$$d_2 = +0.4604$$

and so $X^1 = X^o + d = (2,1) + (-0.92079, +0.4604)$

$= (1.07921,\ 1.4604)$

At this new point, X^1, we compute a new QP approximating problem.

$$\nabla f(X^1) = \begin{bmatrix} 1.78475 \\ -2.17750 \end{bmatrix},\ \nabla^2 f(X^1) = \begin{bmatrix} 6.4595 & -4.4044 \\ -4.4044 & 4.1579 \end{bmatrix}$$

$$h(X^1) = -0.42393,\ \nabla h(X^1) = \begin{bmatrix} 1.4604 \\ 1.07921 \end{bmatrix},\ \nabla^2 h(X^1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$g(X^1) = -1.53961 < 0,\ \nabla g(X^1) = \begin{bmatrix} -1 \\ -1 \end{bmatrix},\ \nabla^2 g(X^\circ) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Minimize $\left[1.78475, -2.17750\right]\begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$

$$+ \tfrac{1}{2}\left[d_1\ d_2\right]\begin{bmatrix} 6.4595 & -4.4044 \\ -4.4044 & 4.1579 \end{bmatrix}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

subject to

$$\left[1.4604, 1.07921\right]\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 0.42393$$

$$\left[-1,\ -1\right]\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \le 1.53961$$

SQP algorithm

Step 0: Select an estimate X^o of the optimal solution, and let t=0. (X^o need not be feasible!)
Step 1: Approximate the problem with a linearly constrained QP problem at X^t.
Step 2: Solve for the optimal d^t.
Step 3: If $d^t \approx 0$, stop;
else, let $X^{t+1} = X^t + d^t$
Increment t and return to step 1.

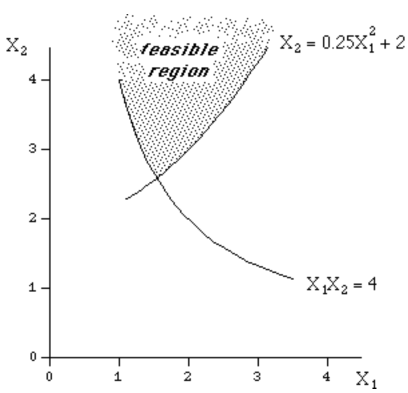
Example

Minimize $\left(X_2 - X_1^2\right)^2 + \left(1 - X_1\right)^2$

subject to

$X_1X_2 \ge 4$,

$X_2 \ge 0.25X_1^2 + 2$



$$\boxed{f(X) = \left(X_2 - X_1^2\right)^2 + \left(1 - X_1\right)^2}$$

Objective

```
Z←F X
⍝
⍝     Objective fn for Successive QP Example
⍝
Z←((X[2]-X[1]*2)*2)+(1-X[1])*2
```

$$\boxed{\begin{aligned} &X_1X_2 \ge 4, \\ &X_2 \ge 0.25X_1^2 + 2 \end{aligned}} \quad \text{i.e.,} \quad \boxed{\begin{aligned} &g_1(X) = 4 - X_1X_2 \le 0 \\ &g_2(X) = 2 + 0.25X_1^2 - X_2 \le 0 \end{aligned}}$$

Inequality Constraints G(x)≤0

```
z←G X
⍝
⍝          Constraint functions for SQP example
⍝
z←4-X[1]×X[2]
z←z,2+(.25×X[1]*2)-X[2]
```

Gradient of objective

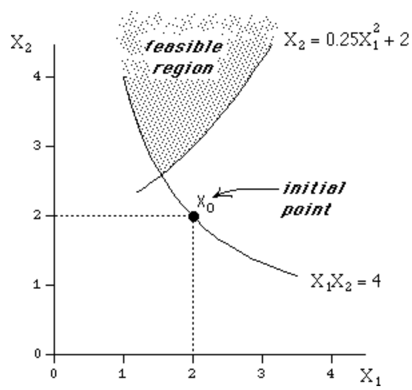
```
G←GRADIENT X
⍝
⍝        Gradient for objective function of SQP example
⍝
G←(4×X[1]*3)+(¯4×X[1]×X[2])+(2×X[1])-2
G←G,2×(X[2]-X[1]*2)
```

Hessian of objective

```
H←HESSIAN∆F X
⍝
⍝       Hessian function for Objective
⍝
H←2 2⍴0
H[1;1]←(12×X[1]*2)+(¯4×X[2])+2
H[2;2]←2
H[1;2]←H[2;1]←¯4×X[1]
```

Jacobian of Inequality Constraints

```
J←JACOBIAN X
⍝
⍝       Jacobian matrix of inequality constraints
⍝            for SQP example
⍝
J←2 2⍴(-X[2]),(-X[1]),(.5×X[1]),¯1
```



QP Approximation

Hessian of Objective Fn

$$\begin{matrix} 42 & ^{-}8 \\ ^{-}8 & 2 \end{matrix}$$

Linear Terms of Objective

```
   i:    1  2
C[i]: 18 ¯4
```

Jacobian Matrix of Constraints & RHS

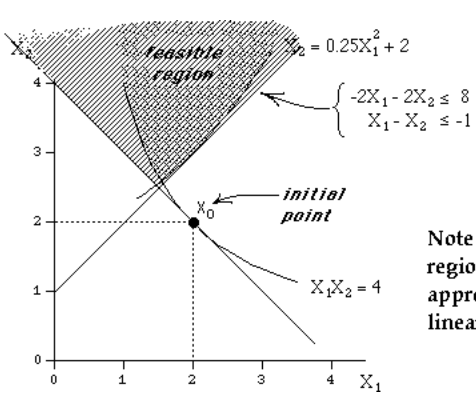
$$\left.\begin{matrix} ^{-}2 & ^{-}2 \\ 1 & ^{-}1 \end{matrix}\;\right|\; \begin{matrix} 0 \\ -1 \end{matrix}$$

Minimize $21d_1^2 - 8d_1d_2 + d_2^2 + 18d_1 - 4d_2$

subject to

$-2d_1 - 2d_2 \le 0$

$d_1 - d_2 \le -1$



Note that feasible region of the QP approximation is linear!

Current SQP Parameter Settings

Convergence criteria:

(The algorithm terminates when either of the following is satisfied, where |Δx| is the change in optimal x between two successive QP problems, and ΔF(x) is the change in the objective function.)

$$\max|\Delta x| \le 0.001$$
$$|\Delta F(x)| \le 0.001$$

Iteration # 1

```
X = 2 2
F(x) = 5
∇F(x) = 18 ¯4
```

∇∇F(x) (Hessian matrix)

$$\begin{matrix} 42 & ^{-}8 \\ ^{-}8 & 2 \end{matrix}$$

```
G(x) = 0 1
```

∇G(x) (Jacobian matrix)

$$\begin{matrix} ^{-}2 & ^{-}2 \\ 1 & ^{-}1 \end{matrix}$$

Feasible region of subproblem (in terms of X_1 & X_2)

$$\begin{cases} -2d_1 - 2d_2 \le 0 \\ d_1 - d_2 \le -1 \\ X_1 = X_1^0 + d_1 \\ X_2 = X_2^0 + d_2 \end{cases} \Rightarrow \begin{cases} -2X_1 - 2X_2 \le 8 \\ X_1 - X_2 \le -1 \end{cases}$$

$$X_1 = X_1^0 + d_1,\; X_2 = X_2^0 + d_2 \Rightarrow \begin{cases} d_1 = X_1 - X_1^0 \\ d_2 = X_2 - X_2^0 \end{cases}$$

$$\Rightarrow d_1 = X_1 - 2,\; d_2 = X_2 - 2$$

Tableau

(before adding artificial variable)

1	2	3	4	5	6	b
¯2	¯2	0	0	1	0	0
1	¯1	0	0	0	1	¯1
42	¯8	¯2	1	0	0	¯18
¯8	2	¯2	¯1	0	0	4

These represent the K.T. conditions:
- Rows 1 through 2 represent ∇g(x)Δx ≤ -g(x)
- Rows 3 through 4 represent H(x)Δx - ∇g(x)U = -∇f(x)

Variable numbers:
- Δx: 1 2
- Y: 5 6 (slack variables for ∇g(x)Δx ≤ -g(x) constraints)
- U: 3 4 (multipliers for ∇g(x)Δx ≤ -g(x) constraints)

Δx is unrestricted in sign, while Y & U ≥0

TABLEAU

(after pivoting Δx and slack variables into basis)

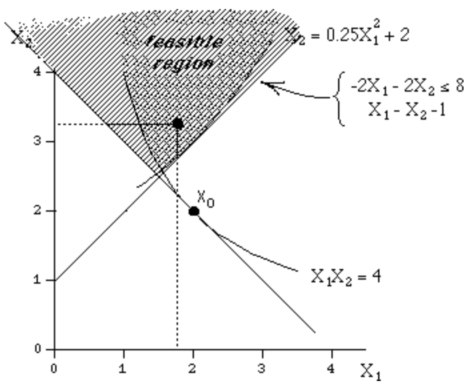
1	2	3	4	5	6	b
0	0	¯12	¯4	1	0	2
0	0	¯4	¯1.4	0	1	0.4
1	0	¯1	¯0.3	0	0	¯0.2
0	1	¯5	¯1.7	0	0	1.2

Because the variable d_1 is not required to be nonnegative, this is a feasible basic solution, and no further pivoting is required!

Optimal QP Subproblem Solution

Primal Variables: Δx = ¯0.2 1.2
Slack: y = 2 0.4
Dual Variables: u = 0 0
QP subproblem objective Function
(approximate improvement ΔF): ¯4.2

X = X + Δx = 1.8 3.2



Iteration # 2

X = 1.8 3.2
F(x) = 0.6416
∇F(x) = 1.888 ¯0.08

∇∇F(x) (Hessian matrix)

28.08	¯7.2
¯7.2	2

G(x) = ¯1.76 ¯0.39

∇G(x) (Jacobian matrix)

¯3.2	¯1.8
0.9	¯1

Lagrange multipliers U = 0 0

QP Approximation

Hessian of Objective Fn

28.08	¯7.2
¯7.2	2

Linear Terms of Objective

i:	1	2
C[i]:	1.888	¯0.08

Jacobian Matrix of Constraints & RHS

¯3.2	¯1.8	1.76
0.9	¯1	0.39

TABLEAU

(after pivoting Δx and slack variables into basis)

1	2	3	4	5	6	b
0	0	¯45.0007	¯13	1	0	¯5.33837
0	0	¯13	¯3.875	0	1	¯1.57
1	0	¯4.48148	¯1.25	0	0	¯0.740741
0	1	¯17.0333	¯5	0	0	¯2.62667

Only the first two rows of the tableau have infeasibility, since there are no nonnegative restrictions on the step vector d.

TABLEAU

(with artificial variable included)

1	2	3	4	5	6	a	b	
0	0	¯45.0007	¯13	1	0	¯1	¯5.33837	← **pivot row**
0	0	¯13	¯3.875	0	1	¯1	¯1.57	
1	0	¯4.48148	¯1.25	0	0	0	¯0.740741	
0	1	¯17.0333	¯5	0	0	0	¯2.62667	

Artificial variable (a) enters the basis, replacing variable 5 whose complement is 3

Artificial variable (a) enters the basis, replacing variable 5 whose complement is 3

1	2	3	4	5	6	a	b
0	0	45.0007	13	¯1	0	1	5.33837
0	0	32.0007	9.125	¯1	1	0	3.76837
1	0	¯4.48148	¯1.25	0	0	0	¯0.740741
0	1	¯17.0333	¯5	0	0	0	¯2.62667

Entering: 3, Leaving: 6 (Pivot in row 2)

Note that the step variables d_1 & d_2 will never leave the basis, because they are not bounded below by zero!

Tableau

1	2	3	4	5	6	a	b
0	0	0	0.168055	0.406241	¯1.40624	1	0.039135
0	0	1	0.28515	¯0.0312493	0.0312493	0	0.117759
1	0	0	0.0278929	¯0.140043	0.140043	0	¯0.213007
0	1	0	¯0.142951	¯0.532279	0.532279	0	¯0.620841

Entering: 4, Leaving: 7 (Pivot in row 1)

artificial variable

Tableau

1	2	3	4	5	6	a	b
0	0	0	1	2.41731	¯8.36776	5.95045	0.232871
0	0	1	0	¯0.720545	2.41731	¯1.69677	0.0513559
1	0	0	0	¯0.207469	0.373444	¯0.165975	¯0.219502
0	1	0	0	¯0.186722	¯0.6639	0.850622	¯0.587552

We now have a basic feasible solution in our tableau!

Optimal QP Subproblem Solution

Primal Variables: Δx = ¯0.219502 ¯0.587552
Slack: y = 0 0
Dual Variables: u = 0.0513559 0.232871
QP subproblem objective Function
(approximate improvement ΔF): ¯0.274311

X = X + Δx = 1.5805 2.61245

Iteration # 3

X = 1.5805 2.61245
F(x) = 0.350082
∇F(x) = 0.437289 0.228949

<u>∇∇F(x) (Hessian matrix)</u>

21.5259 ¯6.32199
¯6.32199 2

G(x) = ¯0.128969 0.0120453

Note that X is infeasible in the second constraint!

<u>∇G(x) (Jacobian matrix)</u>

¯2.61245 ¯1.5805
0.790249 ¯1

Lagrange multipliers U = 0.0513559 0.232871

QP Approximation

<u>Hessian of Objective Fn</u>

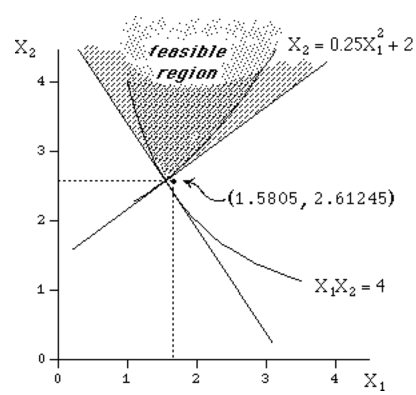
21.5259 ¯6.32199
¯6.32199 2

<u>Linear Terms of Objective</u>

i: 1 2
C[i]: 0.437289 0.228949

<u>Jacobian Matrix of Constraints & RHS</u>

¯2.61245	¯1.5805	0.128969
0.790249	¯1	¯0.0120453



Tableau

(before adding artificial variable)

1	2	3	4	5	6	b
¯2.61245	¯1.5805	0	0	1	0	0.128969
0.790249	¯1	0	0	0	1	¯0.0120453
21.5259	¯6.3219	¯2.61245	0.790249	0	0	¯0.437289
¯6.32199	2	¯1.5805	¯1	0	0	¯0.228949

These represent the K.T. conditions:
Rows 1 through 2 represent $\nabla g(x)\Delta x \le -g(x)$
Rows 3 through 4 represent $H(x)\Delta x - \nabla g(x)U = -\nabla f(x)$

Variable numbers:
Δx: 1 2
Y: 5 6 (slack variables for $\nabla g(x)\Delta x \le -g(x)$ constraints)
U: 3 4 (multipliers for $\nabla g(x)\Delta x \le -g(x)$ constraints)

Δx is unrestricted in sign, while Y & U ≥0

TABLEAU

(after pivoting Δx and slack variables into basis)

1	2	3	4	5	6	b
0	0	¯38.7871	¯12.487	1	0	¯5.78006
0	0	¯12.487	¯4.14466	0	1	¯1.91137
1	0	¯4.93378	¯1.53735	0	0	¯0.752866
0	1	¯16.3859	¯5.35955	0	0	¯2.49428

TABLEAU

(with artificial variable included)

1	2	3	4	5	6	a	b	
0	0	¯38.7871	¯12.487	1	0	¯1	¯5.78006	← pivot row
0	0	¯12.487	¯4.14466	0	1	¯1	¯1.91137	feasible!
1	0	¯4.93378	¯1.53735	0	0	0	¯0.752866	
0	1	¯16.3859	¯5.35955	0	0	0	¯2.49428	

Artificial variable (a) enters the basis, replacing variable 5, whose complement is 3

Tableau

1	2	3	4	5	6	a	b
0	0	38.7871	12.487	¯1	0	1	5.78006
0	0	26.3002	8.34233	¯1	1	0	3.86868
1	0	¯4.93378	¯1.53735	0	0	0	¯0.752866
0	1	¯16.3859	¯5.35955	0	0	0	¯2.49428

Entering: 3, Leaving: 6 (Pivot in row 2)

Tableau

1	2	3	4	5	6	a	b
0	0	0	0.183821	0.474788	¯1.47479	1	0.074569
0	0	1	0.317197	¯0.0380226	0.0380226	0	0.147097
1	0	0	0.0276343	¯0.187595	0.187595	0	¯0.0271193
0	1	0	¯0.161983	¯0.623035	0.623035	0	¯0.0839547

Entering: 4, Leaving: 7 (Pivot in row 1)

Tableau

1	2	3	4	5	6	a	b
0	0	0	1	2.58288	¯8.02294	5.44006	0.40566
0	0	1	0	¯0.857304	2.58288	¯1.72557	0.0184231
1	0	0	0	¯0.258971	0.409303	¯0.150332	¯0.0383294
0	1	0	0	¯0.204652	¯0.676549	0.8812	¯0.0182445

The artificial variable has now left the basis (i.e., has been driven to zero).

Optimal QP Subproblem Solution

Primal Variables: Δx = ¯0.0383294 ¯0.0182445
Slack: y = 0 0
Dual Variables: u = 0.0184231 0.40566
QP subproblem objective Function
(approximate improvement ΔF): ¯0.00921389

X = X + Δx = 1.54217 2.5942

Iteration # 4

X = 1.54217 2.5942
F(x) = 0.340568
$\nabla F(x)$ = ¯0.247602 0.43184

$\nabla\nabla F(x)$ (Hessian matrix)

20.1626	¯6.16867
¯6.16867	2

G(x) = ¯0.000699299 0.000367285

$\nabla G(x)$ (Jacobian matrix)

¯2.5942	¯1.54217
0.771084	¯1

Lagrange multipliers U = 0.0184231 0.40566

QP Approximation

Hessian of Objective Fn

20.1626	¯6.16867
¯6.16867	2

Linear Terms of Objective

i:	1	2
C[i]:	¯0.247602	0.43184

Jacobian Matrix of Constraints & RHS

¯2.5942	¯1.54217	0.000699299
0.771084	¯1	¯0.000367285

Tableau

(before adding artificial variable)

1	2	3	4	5	6	b
¯2.5942	¯1.54217	0	0	1	0	0.000699299
0.771084	¯1	0	0	0	1	¯0.000367285
20.1626	¯6.16867	¯2.5942	0.771084	0	0	0.247602
¯6.16867	2	¯1.54217	¯1	0	0	¯0.43184

These represent the K.T. conditions:
Rows 1 through 2 represent $\nabla g(x)\Delta x \le -g(x)$
Rows 3 through 4 represent $H(x)\Delta x - \nabla g(x)U = -\nabla f(x)$

Variable numbers:
Δx: 1 2
Y: 5 6 (slack variables for $\nabla g(x)\Delta x \le -g(x)$ constraints)
U: 3 4 (multipliers for $\nabla g(x)\Delta x \le -g(x)$ constraints)

Δx is unrestricted in sign, while Y & U ≥ 0

TABLEAU

(after pivoting Δx and slack variables into basis)

1	2	3	4	5	6	b
0	0	-48.7407	-15.7353	1	0	-7.34678
0	0	-15.7353	-5.20918	0	1	-2.42372
1	0	-6.46892	-2.03574	0	0	-0.954253
0	1	-20.7234	-6.77891	0	0	-3.15916

TABLEAU

(with artificial variable included)

1	2	3	4	5	6	a	b	
0	0	-48.7407	-15.7353	1	0	-1	-7.34678	← pivot row
0	0	-15.7353	-5.20918	0	1	-1	-2.42372	feasible!
1	0	-6.46892	-2.03574	0	0	0	-0.954253	
0	1	-20.7234	-6.77891	0	0	0	-3.15916	

Artificial variable (a) enters the basis, replacing variable 5, whose complement is 3

Tableau

1	2	3	4	5	6	a	b
0	0	48.7407	15.7353	-1	0	1	7.34678
0	0	33.0054	10.5262	-1	1	0	4.92306
1	0	-6.46892	-2.03574	0	0	0	-0.954253
0	1	-20.7234	-6.77891	0	0	0	-3.15916

Entering: 3, Leaving: 6 (Pivot in row 2)

Tableau

1	2	3	4	5	6	a	b
0	0	0	0.190825	0.476751	-1.47675	1	0.0766404
0	0	1	0.318923	-0.0302981	0.0302981	0	0.149159
1	0	0	0.0273461	-0.195996	0.195996	0	0.0106483
0	1	0	-0.169739	-0.62788	0.62788	0	-0.0680623

Entering: 4, Leaving: 7 (Pivot in row 1)

artificial variable

Optimal QP Subproblem Solution

```
Primal Variables: Δx = -0.000334549 0.00010932
      Slack:      y = 0 0
Dual Variables:   u = 0.0210721 0.401625
QP subproblem objective Function
(approximate improvement ΔF):  0.00013141

X = X + Δx = 1.54183 2.59431

***Convergence criterion satisfied:
                             |Δx|≤ 0.001
                             |ΔF|≤ 0.001

F(x) = 0.3407
G(x) = 3.65728E-8 2.79808E-8
```

slightly infeasible in both constraints!

Because the standard QP problem has linear constraints, we were allowed to use only linear approximations to the nonlinear constraint functions.

By optimizing a quadratic approximation of the Lagrangian function, we can make use of 2nd-derivative information about the nonlinear constraint functions!

QP Approximation of Lagrangian Function

Consider

Minimize $f(x)$

subject to

$$h_k(x) = 0,\ k=1,\cdots K$$

which has the Lagrangian function

$$L(x,\lambda) = f(x) + \sum_{k=1}^{K} \lambda_k h_k(x) = f(x) + \lambda^T h(x)$$

The solution of

$$\boxed{\begin{array}{l}\text{Minimize } f(x) + \sum_{k=1}^{K} \lambda_k h_k(x) \\ \text{subject to} \\ \qquad h_k(x) = 0,\ k=1,\cdots K\end{array}}$$

is clearly a solution *also* of the original problem.

Given a current iterate $(\bar{x},\bar{\lambda})$, we can form a quadratic approximation to the (Lagrangian) objective and linear approximation to the equality constraints.

$$L(\bar{x}+d,\bar{\lambda}) = L(\bar{x},\bar{\lambda}) + \left[\nabla_x L(\bar{x},\bar{\lambda})\right]^T d + \frac{1}{2} d^T \nabla_x^2 L(\bar{x},\bar{\lambda}) d + \cdots$$

$$= \left[f(\bar{x}) + \bar{\lambda}^T h(\bar{x})\right] + \left[\nabla f(\bar{x}) + \bar{\lambda}^T \nabla h(\bar{x})\right]^T d + \frac{1}{2} d^T \left[\nabla^2 f(\bar{x}) + \sum_k \bar{\lambda}_k \nabla^2 h_k(\bar{x})\right] d + \cdots$$

$$\left[f(\bar{x}) + \bar{\lambda}^T h(\bar{x})\right] + \quad \text{Minimum} \left[\nabla f(\bar{x}) + \bar{\lambda}^T \nabla h(\bar{x})\right]^T d + \frac{1}{2} d^T \left[\nabla^2 f(\bar{x}) + \sum_k \bar{\lambda}_k \nabla^2 h_k(\bar{x})\right] d$$

subject to

$$\left[\nabla h_k(x)\right]^T d = -h_k(x),\ k=1,\cdots K$$

Unlike the previous approximating QP problem, this QP problem makes use of information about the second derivatives of the constraint functions!

SQP algorithm

Step 0: Select an estimate X^o of the optimal solution, and let t=0. (X^o need not be feasible!)
Step 1: Approximate the problem with a linearly constrained QP problem at X^t .
Step 2: Solve for the optimal d^t.
Step 3: If $d^t \approx 0$, stop;
else, let $X^{t+1} = X^t + d^t$
Increment t and return to step 1.

Step 0 Select initial x^o and multiplier λ^o vector (e.g., $\lambda^o = 0$); set t=0

Step 1 Compute the approximating QP with $\bar{x} = x^t$ and $\bar{\lambda} = \lambda^t$

$$\text{Minimum} \left[\nabla f(\bar{x}) + \bar{\lambda}^T \nabla h(\bar{x})\right]^T d + \frac{1}{2} d^T \left[\nabla^2 f(\bar{x}) + \sum_k \bar{\lambda}_k \nabla^2 h_k(\bar{x})\right] d$$

subject to

$$\left[\nabla h_k(x)\right]^T d = -h_k(x),\ k=1,\cdots K$$

SQP Algorithm

Step 2 Solve for the optimal d^*, and compute the optimal Lagrange multipliers λ^* of the QP.

Step 3 If $d^* \approx 0$, STOP;
Else, let $x^{t+1} = x^t + d^*$, $\lambda^{t+1} = \lambda^*$
Increment t and return to step 1.