HW # 7 SQP/IP | ME 575 | Jesse Wynn

1.) $\boxed{\text{Iteration 5}}$

At $(x^4)^T = [0.2533, -0.1583]$ $\quad f^4 = 4.6070$

$\nabla f^4 = [-1.2680, -0.4449]^T$, $\quad g = -0.3720$

$\nabla g = [-1.0066, 0.75]^T$

$$\nabla L(x^3, \lambda^4) = \begin{bmatrix} -2.910 \\ -0.2761 \end{bmatrix} - (0.7192)\begin{bmatrix} 0.2132 \\ 0.75 \end{bmatrix} = \begin{bmatrix} -3.0633 \\ -0.8155 \end{bmatrix}$$

$$\nabla L(x^4, \lambda^4) = \begin{bmatrix} -1.2680 \\ -0.4449 \end{bmatrix} - (0.7192)\begin{bmatrix} -1.0066 \\ 0.75 \end{bmatrix} = \begin{bmatrix} -0.5441 \\ -0.9843 \end{bmatrix}$$

$$\gamma^3 = \nabla L(x^4, \lambda^4) - \nabla L(x^3, \lambda^4) = \begin{bmatrix} 2.5193 \\ -0.1688 \end{bmatrix}$$

$$\Delta x^3 = \begin{bmatrix} 0.2533 \\ -0.1583 \end{bmatrix} - \begin{bmatrix} -0.3566 \\ -0.0109 \end{bmatrix} = \begin{bmatrix} 0.6099 \\ -0.1474 \end{bmatrix}$$

BFGS update

$$\nabla^2 L^4 = \begin{bmatrix} 4.1583, & 0.1144 \\ 0.1144, & 1.6184 \end{bmatrix}$$

Use Lagrangian Hessian to approximate our function:

$$f_a = 4.6070 + [-1.268, -0.4449]\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} + \frac{1}{2}[\Delta x_1, \Delta x_2]\begin{bmatrix} 4.1583 & 0.1144 \\ 0.1144 & 1.6184 \end{bmatrix}\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix}$$

$$g_a = -0.3720 + [-1.0066, 0.75]\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} \geq 0$$

Plug into KKT Equations assuming constraint is binding:

$$-1.2680 + 4.1583\,\Delta x_1 + 0.1144\,\Delta x_2 - \lambda(-1.0066) = 0$$

$$-0.4449 + 0.1144\,\Delta x_1 + 1.6184\,\Delta x_2 - \lambda(0.75) = 0$$

$$-0.3720 + -1.0066\,\Delta x_1 + 0.75\,\Delta x_2 = 0$$

$\Rightarrow$ $\Delta x_1 = 0.0990$, $\Delta x_2 = 0.6289$, $\lambda = 0.7791$

Since λ is positive, the constraint was actually binding and we get our new point:

$$x^5 = x^4 + \Delta x = \begin{bmatrix} 0.2533 \\ -0.1583 \end{bmatrix} + \begin{bmatrix} 0.0990 \\ 0.6289 \end{bmatrix} = \begin{bmatrix} 0.3523 \\ 0.4706 \end{bmatrix}$$

Next Page →

→ Now we must check our penalty function...

$$P = f + \sum_{i=1}^{vio} \lambda_i |g_i|$$

$$\Rightarrow P = 4.5395 + 0.7791|-0.0098|$$

$$= 4.5472$$

At this point the penalty function is 4.5472, a decrease from 4.87 and so we take the full step.

$\boxed{\text{Iteration 6}}$

$(x^5)^T = [0.3524, 0.4706]$ $f^5 = 4.5396$

$\nabla f^5 = [-1.7837, 0.6930]^T$, $g = -0.0098$

$\nabla g = [-1.2046, 0.75]^T$

Compute Lagrangian Gradients

$$\nabla L(x^4, \lambda^5) = \begin{bmatrix} -1.2680 \\ -0.4449 \end{bmatrix} - (0.7791)\begin{bmatrix} -1.0066 \\ 0.75 \end{bmatrix} = \begin{bmatrix} -0.8452 \\ 0.1087 \end{bmatrix}$$

$$\nabla L(x^5, \lambda^5) = \begin{bmatrix} -1.7837 \\ 0.6930 \end{bmatrix} - (0.7791)\begin{bmatrix} -1.2046 \\ 0.75 \end{bmatrix} = \begin{bmatrix} -0.4838 \\ -1.0292 \end{bmatrix}$$

$$\gamma^4 = \nabla L(x^5, \lambda^5) - \nabla L(x^4, \lambda^5) = \begin{bmatrix} -0.3614 \\ 1.1379 \end{bmatrix}$$

$$\Delta x^4 = \begin{bmatrix} 0.3524 \\ 0.4706 \end{bmatrix} - \begin{bmatrix} 0.2533 \\ -0.1583 \end{bmatrix} = \longrightarrow \begin{bmatrix} 0.0990 \\ 0.6289 \end{bmatrix}$$

BFGS Update:

$$\nabla^2 L^5 = \begin{bmatrix} 4.0140 & -1.2066 \\ -1.2066 & 1.9993 \end{bmatrix}$$

Use the Lagrangian Hessian to approximate our function:

$$f_a = 4.5396 + [-1.7837, 0.6930]\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} + \frac{1}{2}[\Delta x_1, \Delta x_2]\begin{bmatrix} 4.0140 & -1.2066 \\ -1.2066 & 1.9993 \end{bmatrix}\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix}$$

$$g_a = -0.0098 + [-1.2046, 0.75]\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} \geq 0$$

Next Page ⟶

Plug in to KKT equations assuming constraint is binding:

$$-1.7837 + 4.0140\,\Delta x_1 - 1.2066\,\Delta x_2 - \lambda(-1.2046) = 0$$

$$0.6930 - 1.2066\,\Delta x_1 + 1.9993\,\Delta x_2 - \lambda(0.75) = 0$$

$$-0.0098 - 1.2046\,\Delta x_1 + 0.75\,\Delta x_2 = 0$$

$$\Rightarrow \Delta x_1 = 0.1217, \quad \Delta x_2 = 0.2085, \quad \lambda = 1.2841$$

Since λ is positive, the constraint is binding (as assumed) and we get our new point:

$$x^6 = x^5 + \Delta x = \begin{bmatrix} 0.3524 \\ 0.4706 \end{bmatrix} + \begin{bmatrix} 0.1217 \\ 0.2085 \end{bmatrix} = \begin{bmatrix} 0.4741 \\ 0.6791 \end{bmatrix}$$

Now we check our penalty function...

$$P = f + \sum_{i=1}^{\text{vio}} \lambda_i |g_i|$$

$$\Rightarrow P = 4.4830 + 1.2841\,|-0.0150|$$

$$= 4.5022$$

At this point the penalty function is 4.5022, a decrease from 4.5472 and so we take the full step.

Thus $x^6 = \begin{bmatrix} 0.4741 \\ 0.6791 \end{bmatrix}$

End Iteration 6

Comments: The algorithm takes a path toward the optimum that is outside of feasible space. This is interesting & shows that the SQP algorithm need not be initialized from a feasible spot.

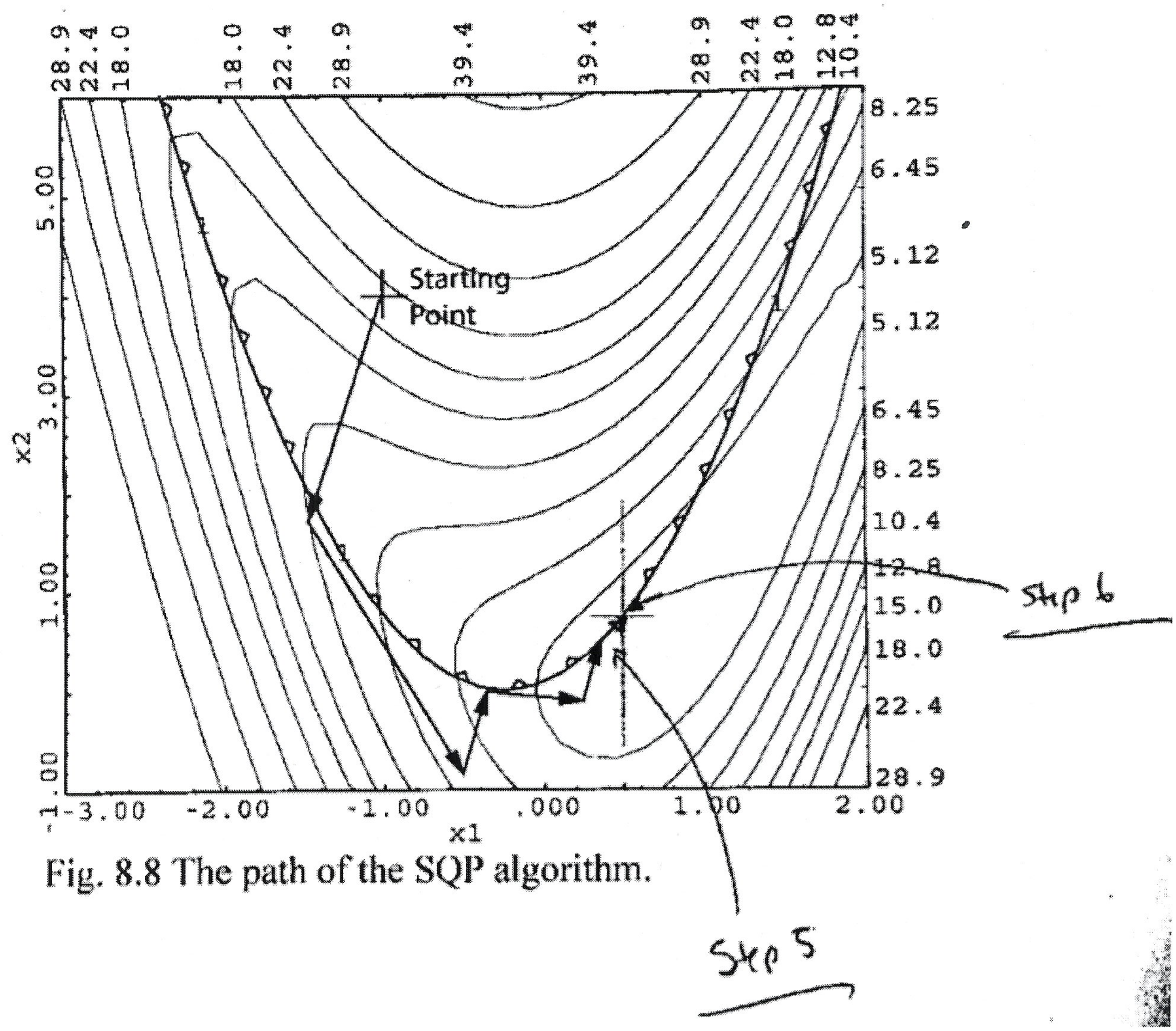

Fig. 8.8 The path of the SQP algorithm.

HW # 7 SQP/IP | ME 575 | Jesse Wynn

Problem 2

IP Iteration 3

$$(x^2)^T = [-0.592, -1.162], \quad s = 0.230, \quad \lambda = 0, \quad f^2 = 8.820$$

$$g = -0.9885$$

$$\nabla f = \begin{bmatrix} -6.7655 \\ -3.0249 \end{bmatrix} \qquad \nabla g = \begin{bmatrix} 0.6840 \\ 0.75 \end{bmatrix}$$

$\Rightarrow$ Update the Hessian of the Lagrangian:

$$\nabla L(x^1, \lambda^2) = \begin{bmatrix} -10.256 \\ -1.435 \end{bmatrix} - (0)\begin{bmatrix} 2.891 \\ 0.75 \end{bmatrix} = \begin{bmatrix} -10.256 \\ -1.435 \end{bmatrix}$$

$$\nabla L(x^2, \lambda^2) = \begin{bmatrix} -6.7655 \\ -3.0249 \end{bmatrix} - (0)\begin{bmatrix} 0.6840 \\ 0.75 \end{bmatrix} = \begin{bmatrix} -6.7655 \\ -3.0249 \end{bmatrix}$$

$$\gamma^1 = \nabla L(x^2, \lambda^2) - \nabla L(x^1, \lambda^2) = \begin{bmatrix} 3.4905 \\ -1.5899 \end{bmatrix}$$

$$\Delta x = \begin{bmatrix} -0.592 \\ -1.162 \end{bmatrix} - \begin{bmatrix} -1.695 \\ 2.157 \end{bmatrix} = \begin{bmatrix} 1.1030 \\ -3.3190 \end{bmatrix}$$

$\Rightarrow$ BFGS Update

$$\Rightarrow \nabla^2 L = \begin{bmatrix} 18.5986 & 5.1292 \\ 5.1292 & 2.1836 \end{bmatrix}$$

Now we evaluate residuals

$$\nabla f^k(x) - \lambda^k \nabla g^k = \begin{bmatrix} -6.7655 \\ -3.0249 \end{bmatrix} - (0)\begin{bmatrix} 0.6840 \\ 0.75 \end{bmatrix} = \begin{bmatrix} -6.7655 \\ -3.0249 \end{bmatrix}$$

$$S^k \Lambda^k e - \mu^k e = [S^k][\lambda^k] - [\mu^k] = [0.230][0] - [0.2] = -0.2$$

$$g^k(x) - s^k = [-0.9885] - [0.230] = -1.2185$$

In matrix form we now have:

$$\begin{bmatrix} 18.5986 & 5.1292 & 0 & -0.6840 \\ 5.1292 & 2.1836 & 0 & -0.75 \\ 0 & 0 & 0 & 0.230 \\ 0.6840 & 0.75 & -1 & 0 \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta s \\ \Delta \lambda \end{bmatrix} = -\begin{bmatrix} -6.7655 \\ -3.0249 \\ -0.2 \\ -1.2185 \end{bmatrix}$$

Solving gives: $\Delta x_1 = -0.1950$, $\Delta x_2 = 2.1419$

$\Delta s = 0.2546$ $\Delta\lambda = 0.8696$.

Check the penalty function

$$P = f^k + v \sum_{i=1}^{viol} |g_i|$$

Our proposed $x^3 = \begin{bmatrix} -0.592 \\ -1.162 \end{bmatrix} + \begin{bmatrix} -0.1950 \\ 2.1419 \end{bmatrix} = \begin{bmatrix} -0.7870 \\ 0.9799 \end{bmatrix}$

$$P = 7.3234 + (-1.2308)\,|0.4466|$$

$$= 6.7737$$

With the full step our merit function decreases from 10.8 to 6.7737, we take the full step. We also take a full step for s and λ since Δs and $\Delta\lambda$ keep $s \geq 0$ and $\lambda \geq 0$

So our new point is:

$(x^3)^T = [-0.7870, 0.9799]$, $s = 0.4846$, $\lambda = 0.8696$,
$g = 0.4466$

Comment: The algorithm starts in feasible space, and then after iterations 1 and 2, it goes into infeasible space. After iteration 3, it is back inside feasible space. By inspection, it looks like IP takes slightly more iterations to reach the optimum.

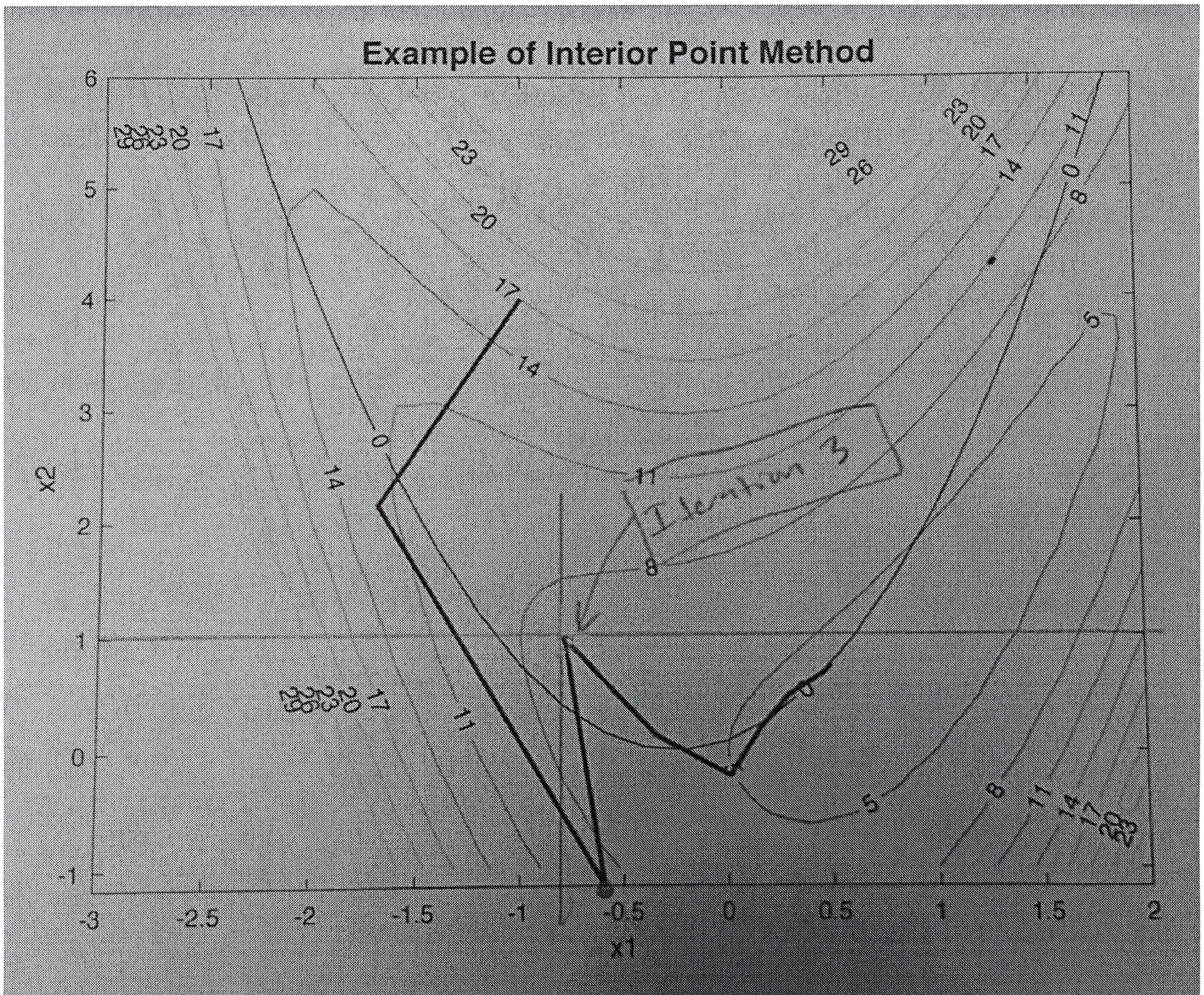

Fig 8.10. The progress of the IP algorithm on Example 1.

```
% clc
% clear

x = [-0.592, -1.162]';

f = get_f(x)

gradf = get_gradf(x)

g = get_g(x)

gradg = get_gradg(x)

H = bfgs(H_old, gamma, delta_x)

x0 = [0, 0, 1, 1];

x = fsolve(@myfun2, x0)

function f = get_f(x)
f = x(1)^4 - 2*x(2)*x(1)^2 + x(2)^2 + x(1)^2 - 2*x(1) + 5;
end

function gradf = get_gradf(x)
gradf = zeros(2,1);
gradf(1) = 4*x(1)^3 - 4*x(2)*x(1) + 2*x(1) -2;
gradf(2) = -2*x(1)^2 + 2*x(2);
end

function g = get_g(x)
g = -(x(1) + 0.25)^2 + 0.75*x(2);
end

function gradg = get_gradg(x)
gradg = zeros(2,1);
gradg(1) = -2*x(1) - 0.5;
gradg(2) = 0.75;
end

% bfgs hessian update
function H_new = bfgs(H_old, gamma, delta_x)

H_new = H_old + ((gamma*gamma')/(gamma'*delta_x)) - ((H_old*delta_x*delta_x'*H_old)/
(delta_x'*H_old*delta_x));
end

function F = myfun(x)
F(1) = -1.7837 + 4.0140*x(1) - 1.2066*x(2) - x(3)*(-1.2046);
F(2) = 0.6930 - 1.2066*x(1) + 1.9993*x(2) - x(3)*(0.75);
F(3) = -0.0098 - 1.2046*x(1) + 0.75*x(2);
end

function F = myfun2(x)
F(1) = 18.5986*x(1) + 5.1292*x(2) -0.6840*x(4) - 6.7655;
F(2) = 5.1292*x(1) + 2.1836*x(2) - 0.75*x(4) - 3.0249;
F(3) = 0.230*x(4) - 0.2;
F(4) = 0.6840*x(1) + 0.75*x(2) - 1*x(3) -1.2185;
end
```

ME 575
HW #7 Due Apr 4, 11:50 p.m.
SQP and IP Algorithms

1. (30 pts) Do the next two iterations of the SQP example problem in the notes (Chap. 8 Section 8.3.8, starting on page 12), picking up where the example leaves off, i.e. do Iterations 5 and 6, starting from $\mathbf{x}^T = [0.2533, -0.1583]$ as given on page 19. If you have to cut the step back because the penalty function increases, cut it back by 0.5.

   Show all calculations. Sketch the resulting steps on Fig. 8.8 taken from the notes. Comment on the path of the algorithm.

2. (20 pts) Do the third iteration of the IP example problem in the notes (Chap. 8, Section 8.4.4, starting on page 26), picking up where the example leaves off, i.e. do Iteration 3, starting from $\mathbf{x}^T = [-0.592, -1.162]$ as given on page 28. If you have to cut the step back because the merit function increases, cut it back by 0.5.

   Show all calculations. Sketch the resulting step on Fig. 8.10 taken from the notes. Comment on the path of the algorithm.