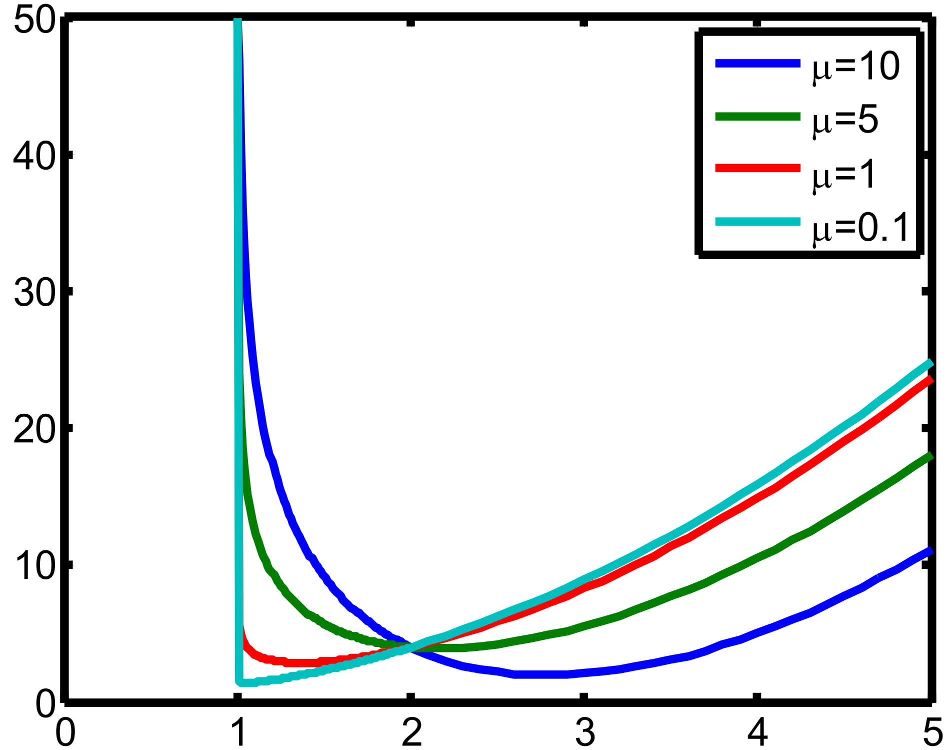
$$\min_{x} x^2 \text{ s.t. } x - 1 \geq 0$$

$$\begin{aligned} \min_{x} \quad & x^2 - \mu \ln s \\ \text{s.t.} \quad & x - 1 = s \\ & s \geq 0 \end{aligned}$$

$$g(x) = x^2 - \mu \ln(x-1)$$



$$x^* = \frac{1 + \sqrt{1 + 2\mu}}{2}$$

μ	x*
10	2.78
5	2.15
1	1.36
0.1	1.04

Lagrangian and KKT system

$$L(x, s, z) = x^2 - \mu \ln(s) - z(x - 1 - s)$$

$$\begin{aligned} L_x &= 2x - z \\ L_s &= -\frac{\mu}{s} + z \\ L_z &= x - 1 - s \end{aligned} \qquad F = \begin{cases} 2x - z = 0 \\ sz - \mu = 0 \\ x - 1 - s = 0 \end{cases}$$

The Jacobian of F is $J = \begin{bmatrix} 2 & 0 & -1 \\ 0 & z & s \\ 1 & -1 & 0 \end{bmatrix}$

Solving the Newton's direction

$$J_k p_k = -F_k, \qquad x_{k+1} = x_k + \alpha_k p_k$$

$$\begin{bmatrix} 2 & 0 & -1 \\ 0 & z & s \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} p_x \\ p_s \\ p_z \end{bmatrix} = - \begin{bmatrix} 2x - z \\ sz - \mu \\ x - 1 - s \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & \sigma & -1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} p_x \\ p_s \\ -p_z \end{bmatrix} = - \begin{bmatrix} 2x - z \\ z - \mu/s \\ x - 1 - s \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & -1/\sigma \end{bmatrix} \begin{bmatrix} p_x \\ -p_z \end{bmatrix} = - \begin{bmatrix} 2x - z \\ x - 1 - \mu/z \end{bmatrix}$$

Singularity of the KKT system

$$F = \begin{cases} 2x - z = 0 \\ sz - \mu = 0 \\ x - 1 - s = 0 \end{cases} \qquad A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & \sigma & -1 \\ 1 & -1 & 0 \end{bmatrix}$$

μ	x*	z	s	σ=z/s	cond(A)
10	2.78	5.66	1.78	3.18	5.50
5	2.15	4.30	1.15	3.74	6.57
1	1.36	2.72	0.36	7.56	14.37
0.1	1.04	2.08	0.04	52	120.8

Modification of the matrix

$$\hat{A} = \begin{bmatrix} 1 & & \\ & s & \\ & & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 \\ 0 & \sigma & -1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & & \\ & s & \\ & & 1 \end{bmatrix}$$

μ	x^*	z	s	$s\sigma s = \mu$	cond($\hat{A}$)
10	2.78	5.66	1.78	10	15.36
5	2.15	4.30	1.15	5	8.52
1	1.36	2.72	0.36	1	4.96
0.1	1.04	2.08	0.04	0.1	23.54

Example of the line search

Using $\mu_k = 0.2 s_k z_k, \alpha_k = 1$

k	x_k	z_k	s_k	μ_k	px_k
0	4	5	2	10	-2.17
1	1.83	3.67	0.83	.306	-0.52
2	1.37	2.64	0.31	.084	-0.23
3	1.08	2.17	.087	.019	-.073
4	1.01	2.03	.014	.003	-.013
5	1.001	2.001	.001	.0003	-.001

Trust region SQP

SQP

chap18

$$\begin{aligned} \min_p \quad & m_k(p) = f_k + \nabla f_k^T p + \tfrac{1}{2} p^T \nabla_{xx}^2 L_k p \\ \text{s.t.} \quad & A_k p + c_k = r_k, \|p\|_2 \le \Delta_k \end{aligned}$$

$$L(x, s, z) = x^2 - \mu \ln(s) - z(x - 1 - s)$$

$$\nabla_{x,s} L = \begin{bmatrix} 2x - z \\ -\mu s^{-1} + z \end{bmatrix}, \nabla_{x,s} f = \begin{bmatrix} 2x \\ -\mu s^{-1} \end{bmatrix}, A_{x,s} = [1, -1]$$

$$L_{xx,ss} = \begin{bmatrix} 2 & 0 \\ 0 & \mu s^{-2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & z s^{-1} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & \sigma \end{bmatrix}$$

$$\begin{aligned} \min_{p_x, p_s} \quad & m(p_x, p_s) = 2x p_x - (\mu/s) p_s + p_x^2 + \sigma p_s^2 / 2 \\ \text{s.t.} \quad & p_x - p_s + (x - 1 - s) = r_I \\ & \|(p_x, p_s)\| \le \Delta \end{aligned}$$

Scaling

- The Hessian of model problem is getting ill-conditioned as s→0 (σ→∞)

$$L_{xx,ss} = \begin{bmatrix} 2 & 0 \\ 0 & zs^{-1} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & \sigma \end{bmatrix}$$

- Let $\hat{p}_s = s^{-1} p_s$. The scaled problem is

$$\begin{aligned} \min_{p_x, p_s} \quad & \hat{m}(p_x, p_s) = 2x p_x - \mu \hat{p}_s + p_x^2 + \sigma s^2 \hat{p}_s^2 / 2 \\ \text{s.t.} \quad & p_x - s\hat{p}_s + (x - 1 - s) = r_I \\ & \|(p_x, \hat{p}_s)\| \leq \Delta \end{aligned}$$

  – Physical meaning of the scaling is to pull the solution s≅0 back to the central path fast.

Solve the dual problem

- The KKT condition: $\nabla f - zA^T = 0$
- But $\begin{bmatrix} 2x \\ -\mu/s \end{bmatrix} = z \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is overdetermined
  - Solved by linear least square (normal equation)

$$z = (AA^T)^{-1} A \begin{bmatrix} 2x \\ -\mu/s \end{bmatrix} = \frac{2x + \mu/s}{2} = x + z/2$$

- For scaled system $\hat{A} = [1, -s]$ $\nabla \hat{f} = \begin{bmatrix} 2x \\ -\mu \end{bmatrix}$

$$\hat{z} = (\hat{A}\hat{A}^T)^{-1} \hat{A} \begin{bmatrix} 2x \\ -\mu \end{bmatrix} = \frac{2x + s\mu}{1 + s^2} \to 2x$$

Example of the trust region SQP

Using $\mu_k = 0.2 s_k z_k, \Delta_k = 1, r_I = 0$

k	x_k	z_k	s_k	μ_k
0	4	5	2	10
1	3.00	5.20	2.00	10
2	2.11	2.92	1.10	2.08
3	1.36	2.61	.364	.648
4	1.02	2.05	.022	.191
5	1.00	2.00	.000	.009