POM-HW4-Jasenv@CC98

Solution 1:

$$\min: q_k(p) = rac{1}{2} p^T B_k p + g_k^T p + f_k$$

By the property of quadratic function, the minimizer is : $p_k = -B_k^{-1}g_k \text{. Consider: } \langle p_k, -g_k \rangle = \left\langle -B_k^{-1}g_k, -g_k \right\rangle = g_k^T B_k^{-T} g_k.$ Since B_k is symmetric positive definite, $\langle p_k, -g_k \rangle = g_k^T B_k^{-T} g_k > 0$, which means that p_k is a decent direction of f at x_k .

Solution 2:

Let:

$$egin{align} F(lpha) &= q_k \left(lpha g_k
ight) = rac{1}{2} \left(g_k^T B_k g_k
ight) lpha^2 + \left(g_k^T g_k
ight) lpha + f_k \ &s.t. \quad \|lpha g_k\| \leqslant \Delta_k, \quad |lpha| \leqslant rac{\Delta_k}{\|g_k\|} \ \end{aligned}$$

equivalently, to minimize $F(\alpha)$ under this constraint. Discuss it in different conditions.

And we can know:

$$egin{aligned} F(lpha)' &= \left(g_k^T B_k g_k
ight) lpha + \left(g_k^T g_k
ight) \ F(lpha_0) &= 0 \Rightarrow lpha_0 = -rac{\left\|g_k
ight\|^2}{g_k^T B_k g_k} \end{aligned}$$

1. $g_k^T B_k g_k \leqslant 0$: $\alpha_0 \geq 0$, $F(\alpha)$ is parabola with opening downward and its axis of symmetry $\alpha = \alpha_0 \geq 0$, thus

$$F_{\min}\left(lpha
ight)=F\left(-rac{\Delta_k}{\|gk\|}
ight)$$
.So, $-rac{\Delta_k}{\|g_k\|}\cdot g_k=lpha g_k=- au_krac{\Delta_k}{\|g_k\|}g_k\Rightarrow au_k=1$

2. $g_k^T B_k g_k > 0$: $\alpha_0 < 0$, $F(\alpha)$ is parabola with opening upward and its axis of symmetry $\alpha = \alpha_0 < 0$

•
$$lpha_0\leqslant rac{-\Delta k}{\|g_k\|}\Rightarrow rac{\|g_k\|^3}{\Delta_k\cdot \left(g_k^TB_kg_k
ight)^2}\geqslant 1$$
, $F(lpha)$ increases in $\left[-rac{\Delta_k}{\|g_k\|},rac{\Delta_k}{\|g_k\|}
ight]$, thus

$$\operatorname{argmin} F(lpha) = -rac{\Delta_k}{\|g_k\|} \Rightarrow au_k = 1.$$

$$ullet$$
 $lpha_0>rac{-\Delta_k}{\|g_k\|}\Rightarrowrac{\|g_k\|^3}{\Delta_k(g_k^TB_kg_k)}<1$, $lpha_0\in\left[-rac{\Delta_k}{\|g_k\|},rac{\Delta_k}{\|g_k\|}
ight]$,thus

$$\operatorname{argmin} F(lpha) = lpha_0 = -rac{\left\|g_k
ight\|^2}{g_k^T B_k g_k}$$

thus,

$$lpha_0 g_k = -rac{\left\|g_k
ight\|^2}{g_k^T B_k g_k} \cdot g_k = - au_k rac{\Delta_k g_k}{\left\|g_k
ight\|} \Rightarrow au_k = rac{\left\|g_k
ight\|^3}{\left(\Delta_k
ight) \cdot \left(g_k^T B_k g_k
ight)}$$

So,

$$au_k = \left\{ egin{aligned} 1, ext{ if } g_k^T B_k g_k & \leq 0 \ \min\left\{rac{\|g_k\|^3}{\Delta_k(g_k^T B_k g_k)}, 1
ight\}, ext{ otherwise.} \end{aligned}
ight.$$

Solution 3:

1. follow the computing process below:

$$egin{aligned} P(\lambda) &= -(B+\lambda I)^{-1}g = -ig(Q\Lambda Q^T + \lambda Iig)^{-1}g = -Q(\Lambda + \lambda I)^{-1}Q^Tg \ &= -(q_1,q_2,\cdots,q_n)\left(\Lambda + \lambda I\right)^{-1}igg(egin{aligned} q_1^Tg \ q_2^Tg \ dots \ q_n^Tg \end{aligned} igg) \ &= -igg(egin{aligned} \frac{q_1^Tg}{\lambda_1 + \lambda}, rac{q_2}{\lambda_2 + \lambda}, \cdots, rac{q_n}{\lambda_n + \lambda} \end{pmatrix} \cdot igg(egin{aligned} q_1^Tg \ q_2^Tg \ dots \ q_n^Tg \end{pmatrix} \ &= -\sum_{i=1}^n rac{q_i^Tg}{\lambda_i + \lambda}q_i \end{aligned}$$

2. since
$$Q$$
 is orthogonal, $\ q_i^Tq_j=\delta_{ij}=egin{cases} 1, & i=j \ 0, & i
eq j \end{cases}$ thus,

$$\|P(\lambda)\|^2 = p(\lambda)^T \cdot p(\lambda) = \left(\sum_{i=1}^n rac{q_i^T g}{\lambda_i + \lambda} q_i^T
ight) \cdot \left(\sum_{j=1}^n rac{q_j^T g}{\lambda_{j+\lambda}} q_j
ight) = \sum_{i=1}^n rac{\left(q_i^T g
ight)^2}{\left(\lambda_i + \lambda
ight)^2}$$

by derivation directly:

$$rac{d\left(\|p(\lambda)\|^2
ight)}{d\lambda} = -2\sum_{i=1}^nrac{\left({q_i}^Tg
ight)^2}{\left(\lambda_i+\lambda
ight)^3}$$

Solution 4:

Let $p = \alpha g + \beta B^{-1}g$ and $u = (\alpha, \beta)^T$.

$$\|H(u):=\|p\|^2=lpha^2\|g\|^2+eta^2ig\|B^{-1}gig\|^2+2lphaeta\left(g,B^{-1}g
ight)=rac{1}{2}u^Tar{B}u^T$$

here

$$ar{B} = 2 egin{bmatrix} \|g\|^2 & \left(B^{-1}g,g
ight) \ \left(B^{-1}g,g
ight) & \left\|B^{-1}g
ight\|^2 \end{bmatrix}$$

And \bar{B} is symmetric positive definite. The constraint becomes:

$$\|p\| \leq \Delta \Rightarrow H(u) \leq \Delta^2$$

Let h(u) := m(p).

$$egin{align} h(u) &= f + \left(lpha g + eta B^{-1} g, g
ight) + rac{1}{2} \left(lpha g + eta B^{-1} g, lpha B g + eta g
ight) \ &= f + lpha \|g\|^2 + eta \left(B^{-1} g, g
ight) + rac{lpha^2}{2} (g, B g) + rac{eta^2}{2} \left(B^{-1} g, g
ight) + lpha eta \|g\|^2 \ &= f + ilde{g}^T u + rac{1}{2} u^T ilde{B} u \end{aligned}$$

here

$$ilde{g} = egin{bmatrix} \|g\|^2 \ eta \left(B^{-1}g,g
ight) \end{bmatrix}, \quad ilde{B} = egin{bmatrix} (Bg,g) & \|g\|^2 \ \|g\|^2 & \left(B^{-1}g,g
ight) \end{bmatrix}$$

Therefore it reduces to the new two-dimensional minimization problem:

$$\min_{u \in R^2} h(u), s.\, t.\, H(u) \leq \Delta^2$$

 \tilde{B} is symmetric positive definite. Let u^* be the minimizer of it with no constraint.

$$u^* = -{ ilde B}^{-1} { ilde g}$$

If $H\left(u^{*}\right) \leq \Delta^{2}$, then u^{*} is the solution. If $H\left(u^{*}\right) > \Delta^{2}$, by Lagrange multiplier method, solve the following problem:

$$\nabla h(u) + \lambda \nabla H(u) = 0, \lambda \geq 0$$

 $(ilde{B} + \lambda ar{B})$ is symmetric positive definite. So we get

$$(ilde{B} + \lambda ar{B})u = - ilde{g} \Rightarrow u = -(ilde{B} + \lambda ar{B})^{-1} ilde{g}$$

Thus with the constrain $H(u) \leq \Delta^2$,

$$\operatorname{argmin} h(u) = egin{cases} u^*, & H\left(u^*
ight) \leq \Delta^2 \ -(ilde{B} + \lambda ar{B})^{-1} ilde{g}, & H\left(u^*
ight) > \Delta^2 \end{cases}$$

 $\lambda \geq 0$ satisfies the following equation:

$$H\left(-(ilde{B}+\lambdaar{B})^{-1} ilde{g}
ight)=\Delta^2$$