POM-HW4-Jasenv@CC98

Solution 1:

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

By the property of quadratic function, the minimizer is :
$p_k = -B_k^{-1} g_k$. 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:

$$F(\alpha) = q_k\left(\alpha g_k\right) = \frac{1}{2}\left(g_k^T B_k g_k\right)\alpha^2 + \left(g_k^T g_k\right)\alpha + f_k$$
$$s.t. \quad \|\alpha g_k\| \leqslant \Delta_k, \quad |\alpha| \leqslant \frac{\Delta_k}{\|g_k\|}$$

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

And we can know:

$$F(\alpha)' = \left(g_k^T B_k g_k\right)\alpha + \left(g_k^T g_k\right)$$
$$F(\alpha_0) = 0 \Rightarrow \alpha_0 = -\frac{\|g_k\|^2}{g_k^T B_k g_k}$$

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

   $F_{\min}(\alpha) = F\left(-\frac{\Delta_k}{\|gk\|}\right)$.So,

   $$-\frac{\Delta_k}{\|g_k\|} \cdot g_k = \alpha g_k = -\tau_k \frac{\Delta_k}{\|g_k\|} g_k \Rightarrow \tau_k = 1$$

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

- $\alpha_0 \leqslant \frac{-\Delta k}{\|g_k\|} \Rightarrow \frac{\|g_k\|^3}{\Delta_k \cdot \left(g_k^T B_k g_k\right)^2} \geqslant 1$, $F(\alpha)$ increases in $\left[-\frac{\Delta_k}{\|g_k\|}, \frac{\Delta_k}{\|g_k\|}\right]$,thus

$$\operatorname{argmin} F(\alpha) = -\frac{\Delta_k}{\|g_k\|} \Rightarrow \tau_k = 1.$$

- $\alpha_0 > \frac{-\Delta_k}{\|g_k\|} \Rightarrow \frac{\|g_k\|^3}{\Delta_k\left(g_k^T B_k g_k\right)} < 1$,$\alpha_0 \in \left[-\frac{\Delta_k}{\|g_k\|}, \frac{\Delta_k}{\|g_k\|}\right]$,thus

$$\operatorname{argmin} F(\alpha) = \alpha_0 = -\frac{\|g_k\|^2}{g_k^T B_k g_k}$$

thus,

$$\alpha_0 g_k = -\frac{\|g_k\|^2}{g_k^T B_k g_k} \cdot g_k = -\tau_k \frac{\Delta_k g_k}{\|g_k\|} \Rightarrow \tau_k = \frac{\|g_k\|^3}{(\Delta_k) \cdot \left(g_k^T B_k g_k\right)}$$

So,

$$\tau_k = \begin{cases} 1, \text{ if } g_k^T B_k g_k \leq 0 \\ \min\left\{\frac{\|g_k\|^3}{\Delta_k\left(g_k^T B_k g_k\right)}, 1\right\}, \text{ otherwise.} \end{cases}$$

Solution 3:

1. follow the computing process below:

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

2. since Q is orthogonal, $q_i^T q_j = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

thus,

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

by derivation directly:

$$\frac{d\left(\|p(\lambda)\|^2\right)}{d\lambda} = -2\sum_{i=1}^{n} \frac{\left({q_i}^T g\right)^2}{\left(\lambda_i + \lambda\right)^3}$$

Solution 4:

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

$$H(u) := \|p\|^2 = \alpha^2 \|g\|^2 + \beta^2 \left\|B^{-1} g\right\|^2 + 2\alpha\beta\left(g, B^{-1} g\right) = \frac{1}{2} u^T \bar{B} u$$

here

$$\bar{B} = 2\begin{bmatrix} \|g\|^2 & \left(B^{-1} g, g\right) \\ \left(B^{-1} g, g\right) & \left\|B^{-1} g\right\|^2 \end{bmatrix}$$

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

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

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

$$\begin{aligned} h(u) &= f + \left(\alpha g + \beta B^{-1} g, g\right) + \frac{1}{2}\left(\alpha g + \beta B^{-1} g, \alpha B g + \beta g\right) \\ &= f + \alpha \|g\|^2 + \beta\left(B^{-1} g, g\right) + \frac{\alpha^2}{2}(g, Bg) + \frac{\beta^2}{2}\left(B^{-1} g, g\right) + \alpha\beta \|g\|^2 \\ &= f + \tilde{g}^T u + \frac{1}{2} u^T \tilde{B} u \end{aligned}$$

here

$$\tilde{g} = \begin{bmatrix} \|g\|^2 \\ \beta\left(B^{-1} g, g\right) \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} (Bg, g) & \|g\|^2 \\ \|g\|^2 & \left(B^{-1} g, g\right) \end{bmatrix}$$

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.

$$u^* = -\tilde{B}^{-1}\tilde{g}$$

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:

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

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

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

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

$$\operatorname{argmin} h(u) = \begin{cases} u^*, & H\left(u^*\right) \leq \Delta^2 \\ -(\tilde{B} + \lambda \bar{B})^{-1}\tilde{g}, & H\left(u^*\right) > \Delta^2 \end{cases}$$

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

$$H\left(-(\tilde{B} + \lambda \bar{B})^{-1}\tilde{g}\right) = \Delta^2$$