

1. The Bregman distance generated by the kernel $\phi(x) = \frac{1}{2t}\|x\|_2^2 - g(x)$ is

$$\begin{aligned} d(x, y) &= \frac{1}{2t}(\|x\|_2^2 - \|y\|_2^2) - g(x) + g(y) - \left(\frac{1}{t}y - \nabla g(y)\right)^T (x - y) \\ &= \frac{1}{2t}(\|x\|_2^2 + \|y\|_2^2) - g(x) + g(y) - \frac{1}{t}x^T y + \nabla g(y)^T (x - y) \\ &= \frac{1}{2t}\|x - y\|_2^2 - g(x) + g(y) - \nabla g(y)^T (x - y). \end{aligned}$$

Substitute y with x_k , the proximal point iteration boils down to

$$\begin{aligned} x_{k+1} &= \arg \min_x g(x) + h(x) + \frac{1}{2t}\|x - x_k\|_2^2 - g(x) + g(x_k) - \nabla g(x_k)^T (x - x_k) \\ &= \arg \min_x th(x) + \frac{1}{2}\|x - x_k\|_2^2 - t\nabla g(x_k)^T (x - x_k) \\ &= \arg \min_x th(x) + \frac{1}{2}\|x - x_k - t\nabla g(x_k)\|_2^2 - \frac{t^2}{2}\|\nabla g(x_k)\|_2^2 \\ &= \text{prox}_{th}(x_k - t\nabla g(x_k)). \end{aligned}$$