

2024/01/16

Exercise 3

Consider the quadratic function

$$J(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T A \mathbf{x} - \mathbf{b}^T \mathbf{x}, \quad (1)$$

where $A \in \mathbb{R}^{n \times n}$ is SPD and $\mathbf{b} \in \mathbb{R}^n$.

1. Compute the gradient and the Hessian of J .
2. Verify that J is strictly convex and find the unique global minimum of J .
3. Let $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^n$ a direction s.t. $\nabla J(\mathbf{x})^T \mathbf{q} < 0$. Compute analytically the step length α that solve the following exact line-search problem

$$\min_{\alpha > 0} J(\mathbf{x} + \alpha \mathbf{q}). \quad (2)$$

$$\nabla J(\mathbf{x}) = A\mathbf{x} - \mathbf{b}$$

$$\nabla^2 J(\mathbf{x}) = A$$

Hessian positive definite \rightarrow function strictly convex. $\because A$ is SPD \rightarrow function is convex

Minimum $\nabla J(\mathbf{x}) = A\mathbf{x} - \mathbf{b} = 0 \rightarrow A\mathbf{x} = \mathbf{b} \quad \mathbf{x}^* = A^{-1}\mathbf{b}$

Exact line search $\min_{\alpha} J(\mathbf{x} + \alpha \mathbf{q})$

define $\phi(\alpha) = J(\mathbf{x} + \alpha \mathbf{q}) = \frac{1}{2} (\mathbf{x} + \alpha \mathbf{q})^T A (\mathbf{x} + \alpha \mathbf{q}) - \mathbf{b}^T (\mathbf{x} + \alpha \mathbf{q})$

$$\begin{aligned} \hookrightarrow \phi(\alpha) &= \frac{1}{2} \mathbf{x}^T A \mathbf{x} + \alpha \mathbf{x}^T A \mathbf{q} + \frac{1}{2} \alpha^2 \mathbf{q}^T A \mathbf{q} - \mathbf{b}^T \mathbf{x} - \alpha \mathbf{b}^T \mathbf{q} \\ &= J(\mathbf{x}) + \alpha (\mathbf{x}^T A \mathbf{q} - \mathbf{b}^T \mathbf{q}) + \frac{1}{2} \alpha^2 \mathbf{q}^T A \mathbf{q} \end{aligned}$$

\hookrightarrow Minimize wrt $\alpha \quad \frac{d\phi}{d\alpha} = 0 = \mathbf{x}^T A \mathbf{q} - \mathbf{b}^T \mathbf{q} + \alpha \mathbf{q}^T A \mathbf{q} = 0$

$$\hookrightarrow \alpha^* = \frac{\mathbf{b}^T \mathbf{q} - \mathbf{x}^T A \mathbf{q}}{\mathbf{q}^T A \mathbf{q}}$$

(using $\nabla J(\mathbf{x}) = A\mathbf{x} - \mathbf{b}$)

$$\begin{aligned} \nabla J(\mathbf{x})^T \mathbf{q} &= (A\mathbf{x} - \mathbf{b})^T \mathbf{q} \\ &= \mathbf{x}^T A \mathbf{q} - \mathbf{b}^T \mathbf{q} \end{aligned}$$

$$\rightarrow \alpha^* = - \frac{\nabla J(\mathbf{x})^T \mathbf{q}}{\mathbf{q}^T A \mathbf{q}}$$