

Scientific Computing HW10

Ryan Chen

November 6, 2024

- Exercise 3.5 of [NW] as a lemma: If B is nonsingular then

$$\|Bx\| \geq \frac{\|x\|}{\|B^{-1}\|}$$

Proof. First $\|x\| = \|B^{-1}Bx\| \leq \|B^{-1}\|\|Bx\|$, then divide by $\|B^{-1}\|$. □

The direction is given by

$$p = -B^{-1}\nabla f(x)$$

We can rewrite this as

$$\nabla f(x) = -Bp$$

Given SPD B we can define $B^{1/2}$ and $B^{-1/2}$, moreover $\|B^{1/2}\| = \|B\|^{1/2}$ and $\|B^{-1/2}\| = \|B\|^{-1/2}$. Then compute

$$\begin{aligned} \cos \theta &= -\frac{\nabla f(x)^T p}{\|\nabla f(x)\| \|p\|} \\ &= \frac{p^T Bp}{\|Bp\| \|p\|} \\ &\geq \frac{p^T Bp}{\|B\| \|p\|^2} && \|Bp\| \leq \|B\| \|p\| \\ &= \frac{\|B^{1/2}p\|^2}{\|B\| \|p\|^2} && p^T Bp = p^T B^{1/2} B^{1/2} p = \|B^{1/2}p\|^2 \\ &= \frac{\|p\|^2}{\|B^{-1/2}\|^2 \|B\| \|p\|^2} && \text{by lemma} \\ &= \frac{1}{\|B^{-1}\| \|B\|} \\ &= \frac{1}{\kappa(B)} \\ &> 0 \end{aligned}$$

This establishes $\cos \theta$ as bounded away from zero.

2. (a) Compute

$$\begin{aligned}
y_k^T s_k &= (\nabla f_{k+1}^T - \nabla f_k^T) \alpha_k p_k \\
&= \alpha_k (\nabla f_{k+1}^T p_k - \nabla f_k^T p_k) \\
&\geq \alpha_k (c_2 \nabla f_k^T p_k - \nabla f_k^T p_k) && \text{Wolfe condition 2} \\
&= \alpha_k (c_2 - 1) \nabla f_k^T p_k \\
&> 0 && \nabla f_k^T p_k \text{ by Problem 1, and } c_2 < 1
\end{aligned}$$

(b) The inner product and norm induced by B_k are

$$(u, v)_{B_k} := v^T B_k u, \quad \|u\|_{B_k} := \sqrt{(u, u)_{B_k}}$$

For all $z \neq 0$,

$$\begin{aligned}
z^T B_{k+1} z &= \|z\|_{B_k}^2 - \frac{(z, s_k)_{B_k}^2}{\|s_k\|_{B_k}^2} + \frac{(z^T y_k)^2}{y_k^T s_k} \\
&> \|z\|_{B_k}^2 - \frac{(z, s_k)_{B_k}^2}{\|s_k\|_{B_k}^2} && y_k^T s_k > 0 \\
&\geq \|z\|_{B_k}^2 - \frac{\|z\|_{B_k}^2 \|s_k\|_{B_k}^2}{\|s_k\|_{B_k}^2} && \text{Cauchy-Schwarz} \\
&= 0
\end{aligned}$$

Thus B_{k+1} is SPD.

3. Code: <https://github.com/RokettoJanpu/scientific-computing-1-redux/blob/main/hw10/hw10p3.ipynb>

4. (a) Compute

$$\partial_x f = 200(y - x^2)(-2x) + 2(1 - x)(-1) = -400x(y - x^2) + 2(x - 1)$$

$$\partial_y f = 200(y - x^2)$$

Critical points are found by solving $\partial_x f = \partial_y f = 0$. From $\partial_y f = 0$ we get $y = x^2$. Plugging this into $\partial_x f = 0$ gives $x = 1$. Thus the only critical point is $(1, 1)$.

The Hessian is

$$H(x, y) = \begin{bmatrix} \partial_x^2 f & \partial_x \partial_y f \\ \partial_x \partial_y f & \partial_y^2 f \end{bmatrix} = \begin{bmatrix} -400(y - 3x^2) + 2 & -400x \\ -400x & 200 \end{bmatrix}$$

Evaluate it at the critical point.

$$H(1, 1) = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}$$

Its eigenvalues λ_1, λ_2 are given by

$$\lambda_1 + \lambda_2 = \text{tr } H(1, 1) = 1002$$

$$\lambda_1 \lambda_2 = \det H(1, 1) = 400$$

From $\lambda_1 + \lambda_2 > 0$ and $\lambda_1 \lambda_2 > 0$ we have $\lambda_1, \lambda_2 > 0$ hence $H(1, 1)$ is SPD. We conclude that the only critical point $(1, 1)$ is a local minimizer.

(b)