MAST30013 – Techniques in Operations Research

Semester 1, 2021

Tutorial 4 Solutions

1. (a) We have that

$$\begin{aligned} \nabla f(x_1, x_2)^T &= \begin{pmatrix} 2x_1 - x_2 + 1, & -x_1 + 3x_2^2 - 1 \end{pmatrix} \\ \nabla^2 f(x_1, x_2) &= \begin{pmatrix} 2 & -1 \\ -1 & 6x_2 \end{pmatrix}. \end{aligned}$$

For $\boldsymbol{x}^* = \left(-\frac{1}{4}, \frac{1}{2}\right)^T$,

$$\begin{aligned} \nabla f\left(-\tfrac{1}{4}, \tfrac{1}{2}\right)^T &= \begin{pmatrix} 0, & 0 \end{pmatrix} \\ \nabla^2 f\left(-\tfrac{1}{4}, \tfrac{1}{2}\right) &= \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}. \end{aligned}$$

Solving

$$\begin{aligned} \det \begin{pmatrix} \lambda - 2 & -1 \\ -1 & \lambda - 3 \end{pmatrix} &= 0 \\ \implies \lambda^2 - 5\lambda + 5 &= 0, \end{aligned}$$

gives eigenvalues $\lambda = \dfrac{5 \pm \sqrt{5}}{2}$. Since they are both positive, the matrix is positive definite, and the stationary point is a local minimum.
For $\boldsymbol{x}^* = \left(-\frac{2}{3}, -\frac{1}{3}\right)^T$,

$$\begin{aligned} \nabla f\left(-\tfrac{2}{3}, -\tfrac{1}{3}\right)^T &= \begin{pmatrix} 0, & 0 \end{pmatrix} \\ \nabla^2 f\left(-\tfrac{2}{3}, -\tfrac{1}{3}\right) &= \begin{pmatrix} 2 & -1 \\ -1 & -2 \end{pmatrix}. \end{aligned}$$

Solving

$$\begin{aligned} \det \begin{pmatrix} \lambda - 2 & -1 \\ -1 & \lambda + 2 \end{pmatrix} &= 0 \\ \implies \lambda^2 - 5 &= 0, \end{aligned}$$

gives eigenvalues $\lambda = \pm\sqrt{5}$. Since one is positive, the other negative, the matrix is neither positive definite nor negative definite, and the stationary point is a saddle point.

(b) The Steepest Descent Method

Step 0.1 $k = 0$, $\boldsymbol{x}^0 = (0, 0)^T$.

Step 0.2 $\boldsymbol{d}^0 = -\boldsymbol{\nabla} f(0, 0) = (-1, 1)^T$, $\| (-1, 1)^T \| = 1.414 > 0.01$.

Step 0.3 $q(t) = f(-t, t) = t^3 + 2t^2 - 2t$, which is minimized when $3t^2 + 4t - 2 = 0$ $\Longrightarrow t = 0.3874$.

Step 0.4 $k = 1$, $\boldsymbol{x}^1 = (0, 0)^T + 0.3874(-1, 1)^T = (-0.3874, 0.3874)^T$.

Step 1.2 $\boldsymbol{d}^1 = -\boldsymbol{\nabla} f(-0.3874, 0.3874) = (0.1623, 0.1623)^T$, $\| (0.1623, 0.1623)^T \| = 0.2295 > 0.01$.

Step 1.3 $q(t) = f(-0.3874 + 0.1623t, 0.3874 + 0.1623t) = 0.0043t^3 + 0.0306t^2 - 0.0527t - 0.4165$, which is minimized when $0.0128t^2 + 0.0612t - 0.0527 = 0$ $\Longrightarrow t = 0.7444$.

Step 1.4 $k = 2$, $\boldsymbol{x}^2 = (-0.3874, 0.3874)^T + 0.7444(0.1623, 0.1623)^T = (-0.2666, 0.5082)^T$.

Step 2.2 $\boldsymbol{d}^2 = -\boldsymbol{\nabla} f(-0.2666, 0.5082) = (0.0415, -0.0415)^T$, $\| (0.0415, -0.0415)^T \| = 0.0587 > 0.01$.

Step 2.3 $q(t) = f(-0.2666 + 0.0415t, 0.5082 - 0.0415t) = -0.0001t^3 + 0.0061t^2 - 0.0034t - 0.4370$, which is minimized when $-0.0002t^2 + 0.0121t - 0.0034 = 0$ $\Longrightarrow t = 0.2851$.

Step 2.4 $k = 3$, $\boldsymbol{x}^3 = (-0.2666, 0.5082)^T + 0.2851(0.0415, -0.0415)^T = (-0.2548, 0.4964)^T$.

Step 3.2 $\boldsymbol{d}^3 = -\boldsymbol{\nabla} f(-0.2548, 0.4964) = (0.0060, 0.0060)^T$, $\| (0.0060, 0.0060)^T \| = 0.0085 < 0.01$.

We have that $\boldsymbol{x}_{\min} = (-0.2548, 0.4964)^T$.

2. In the method of steepest descent we need to minimize

$$q(t) \;=\; f\left(\boldsymbol{x}^k + t\boldsymbol{d}^k\right),$$

for $t > 0$. This is achieved when

$$\frac{dq}{dt} \;=\; \boldsymbol{\nabla} f\left(\boldsymbol{x}^k + t\boldsymbol{d}^k\right)^T \cdot \boldsymbol{d}^k \;=\; 0,$$

using the chain rule for functions of several variables.

Thus, $\boldsymbol{d}^{k+1} = -\boldsymbol{\nabla} f\left(\boldsymbol{x}^{k+1}\right) = -\boldsymbol{\nabla} f\left(\boldsymbol{x}^k + t\boldsymbol{d}^k\right)$ is perpendicular to $\boldsymbol{d}^k$.

3. (a) We want to show that we can find a constant $c \in (0,1)$ such that $||x^{k+1} - x^*|| \leq c||x^k - x^*||$. Now,

$$||x^k - 0|| = \frac{2k}{4^k + k^4 + 1},$$

and

$$||x^{k+1} - 0|| = \frac{2(k+1)}{4^{k+1} + (k+1)^4 + 1}.$$

So the ratio is

$$\begin{aligned} \frac{\| x^{k+1} \|}{\| x^k \|} &= \frac{2(k+1)(4^k + k^4 + 1)}{2k(4^{k+1} + (k+1)^4 + 1)} \\ &= \frac{4^k k + k^5 + k + 4^k + k^4 + 1}{4^{k+1}k + k(k+1)^4 + k}. \end{aligned}$$

If we divide both the numerator and denominator by $4^k k$, then we can clearly see that this term tends to $\frac{1}{4}$. Since we can choose $\frac{1}{4} < c < 1$, the sequence converges linearly.

(b) We want to find the ratio of the terms $\| x^{k+1} - 1 \|$ and $\| x^k - 1 \|$. First we observe that

$$\begin{aligned} ||x^k - 1|| &= \left| \frac{2k^2 - 3k + 8}{2k^2 + 7k - 2} - 1 \right| \\ &= \left| \frac{-10k + 10}{2k^2 + 7k - 2} \right|. \end{aligned}$$

Then

$$\begin{aligned} \frac{||x^{k+1} - 1||}{||x^k - 1||} &= \left| \frac{\frac{-10(k+1) + 10}{2(k+1)^2 + 7(k+1) - 2}}{\frac{-10k + 10}{2k^2 + 7k - 2}} \right| \\ &= \left| \frac{(2k^2 + 7k - 2)(-10k)}{(2k^2 + 11k + 7)(-10k + 10)} \right|, \end{aligned}$$

which converges to 1 as $k \to \infty$. Since we cannot find a $c \in (0,1)$ so that $||x^{k+1} - x^*|| \leq c||x^k - x^*||$ for all sufficiently large k, this sequence converges slower than linearly.