37242 Introduction to Optimisation

Tutorial 9

1. (a) There is one constraint, so we need one element to be in the basis — I'll choose x_1. So the $\mathbf{B}$ matrix is a 1×1 matrix (a scalar), and in fact equals 2. So $\mathbf{B}^{-1} = \frac{1}{2}$. Then

$$\begin{aligned}\mathbf{Z} &= \begin{bmatrix} -\mathbf{B}^{-1}\mathbf{N} \\ \mathbf{I} \end{bmatrix} \\ &= \begin{bmatrix} -\frac{5}{2} & -\frac{1}{2} \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\end{aligned}$$

Then taking the derivative of the objective function:

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2x_1 + 2x_1x_3^2 + 2x_2 \\ 2x_1 + 4x_2^3 + 8 \\ 2x_1^2x_3 \end{bmatrix}.$$

Checking each of the suggested points in turn:

i. Is not feasible, therefore cannot give a stationary point.

ii. Is feasible,

$$\nabla f\left(\begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 8 \\ 0 \end{bmatrix}.$$

Then

$$\begin{aligned}\mathbf{Z}^T\nabla f\left(\begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}\right) &= \begin{bmatrix} -\frac{5}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 8 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 8 \\ 0 \end{bmatrix}\end{aligned}$$

Hence, point (ii) is not a stationary point.

iii. Is feasible,

$$\nabla f\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 4 \\ 10 \\ 2 \end{bmatrix}.$$

Then

$$\begin{aligned}
\mathbf{Z}^T \boldsymbol{\nabla} f \left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right) &= \begin{bmatrix} -\frac{5}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 10 \\ 2 \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\end{aligned}$$

Hence, point (iii) is a stationary point.

(b) For point (iii), we need to consider $\mathbf{Z}^T \boldsymbol{\nabla}^2 f \mathbf{Z}$ (the Hessian of the reduced function). First we calculate the Hessian of the objective function:

$$\boldsymbol{\nabla}^2 f(\mathbf{x}) = \begin{bmatrix} 2 + 2x_3^2 & 2 & 4x_1 x_3 \\ 2 & 12x_2^2 & 0 \\ 4x_1 x_3 & 0 & 2x_1^2 \end{bmatrix}$$

Thus,

$$\boldsymbol{\nabla}^2 f \left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} 4 & 2 & 4 \\ 2 & 0 & 0 \\ 4 & 0 & 2 \end{bmatrix}$$

$$\begin{aligned}
\mathbf{Z}^T \boldsymbol{\nabla}^2 f \left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right) \mathbf{Z} &= \begin{bmatrix} -\frac{5}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 & 4 \\ 2 & 0 & 0 \\ 4 & 0 & 2 \end{bmatrix} \begin{bmatrix} -\frac{5}{2} & -\frac{1}{2} \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} -8 & -5 & -10 \\ 2 & -1 & 0 \end{bmatrix} \begin{bmatrix} -\frac{5}{2} & -\frac{1}{2} \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 15 & -6 \\ -6 & -1 \end{bmatrix}
\end{aligned}$$

We then need to check the eigenvalues of this matrix:

$$\begin{aligned}
\det \begin{bmatrix} 15 - \lambda & -6 \\ -6 & -1 - \lambda \end{bmatrix} &= 0 \\
\text{so } (15 - \lambda)(-1 - \lambda) - (-6)^2 &= 0 \\
\text{so } \lambda^2 - 14\lambda - 51 &= 0
\end{aligned}$$

Using the quadratic formula:

$$\begin{aligned}\lambda &= \frac{14 \pm \sqrt{196 - 4 \times -51}}{2} \\ &= -3 \text{ or } 17.\end{aligned}$$

Since these values are not both positive, the matrix is not positive definite. Hence point (iii) is not a local minimum (it's a type of saddle point).

2. (a) The Lagrangian

$$\begin{aligned}L(\mathbf{x}, \boldsymbol{\mu}) &= x_1^2 + 3x_1x_2 + 9x_2^2 + x_3^2 \\ &\quad -\mu_1(x_1 - x_2 + x_3 - 4) - \mu_2(2x_1 + x_2 + 5x_3 - 8).\end{aligned}$$

(b)

$$\boldsymbol{\nabla}_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\mu}) = \begin{bmatrix} 2x_1 + 3x_2 - \mu_1 - 2\mu_2 \\ 3x_1 + 18x_2 + \mu_1 - \mu_2 \\ 2x_3 - \mu_1 - 5\mu_2 \end{bmatrix}$$

$$\boldsymbol{\nabla}_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\mu}) = \begin{bmatrix} 3.1 - \mu_1 - 2\mu_2 \\ -4.8 + \mu_1 - \mu_2 \\ 1.4 - \mu_1 - 5\mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Adding the first two rows gives $-1.7 - 3\mu_2 = 0$, so $\mu_2^* \approx -0.5667$. Substituting this into row 1 gives $\mu_1^* \approx 4.2333$. We can see by substitution that this works for all three rows.

$$\boldsymbol{\nabla}_{\boldsymbol{\mu}} L(\mathbf{x}^*, \boldsymbol{\mu}) = \begin{bmatrix} -(x_1^* - x_2^* + x_3^* - 4) \\ -(2x_1^* + x_2^* + 5x_3^* - 8) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

so $\boldsymbol{\nabla} L(\mathbf{x}^*, \boldsymbol{\mu}^*) = \mathbf{0}$, which confirms that $\mathbf{x}^*$ is a stationary point for the constrained optimisation problem.

3. (a)

$$L(\mathbf{x}, \mu) = x_1^2 + x_1x_2 + x_2^2 + 4x_1 - \mu(x_1 + 3x_2 - 1).$$

(b) The derivative of the Lagrangian is

$$\boldsymbol{\nabla} L(\mathbf{x}, \mu) = \begin{bmatrix} 2x_1 + x_2 + 4 - \mu \\ x_1 + 2x_2 - 3\mu \\ -(x_1 + 3x_2 - 1) \end{bmatrix}.$$

We wish to find where this vector is $\mathbf{0}$, that is to solve

$$\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & -3 \\ 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \mu \end{bmatrix} = \begin{bmatrix} -4 \\ 0 \\ 1 \end{bmatrix}$$

Gaussian elimination gives:

$$\begin{aligned}
&\left[\begin{array}{ccc|c} 2 & 1 & -1 & -4 \\ 1 & 2 & -3 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right] \\
\sim &\left[\begin{array}{ccc|c} 2 & 1 & -1 & -4 \\ 0 & \frac{3}{2} & -\frac{5}{2} & 2 \\ 0 & \frac{5}{2} & \frac{1}{2} & 3 \end{array}\right] \qquad \begin{array}{l} R_2 \leftarrow R_2 - \frac{1}{2}R_1 \\ R_3 \leftarrow R_3 - \frac{1}{2}R_1 \end{array} \\
\sim &\left[\begin{array}{ccc|c} 2 & 1 & -1 & -4 \\ 0 & 1 & -\frac{5}{3} & \frac{4}{3} \\ 0 & 0 & \frac{14}{3} & -\frac{1}{3} \end{array}\right] \qquad \begin{array}{l} R_2 \leftarrow \frac{2}{3}R_2 \\ R_3 \leftarrow R_3 - \frac{5}{2} \text{ new } R_2 \end{array}
\end{aligned}$$

So, by backwards substitution, $\mu = -\frac{1}{14}, x_2 = \frac{17}{14}, x_1 = -\frac{37}{14}$. This is the only solution, so the only stationary point is at

$$\mathbf{x} = \begin{bmatrix} -\frac{37}{14} \\ \frac{17}{14} \end{bmatrix}$$

4. Consider the problem

$$\begin{array}{ll} \min & f(x_1, x_2, x_3) = 3x_1^2 - \frac{1}{2}x_2^2 - \frac{1}{2}x_3^2 + x_1x_2 - x_1x_3 + 2x_2x_3 \\ \text{s.t.} & 2x_1 - x_2 + x_3 = 2. \end{array}$$

(a) The Lagrangian is

$$L(x_1, x_2, x_3, \mu) = 3x_1^2 - \frac{1}{2}x_2^2 - \frac{1}{2}x_3^2 + x_1x_2 - x_1x_3 + 2x_2x_3 - \mu(2x_1 - x_2 + x_3 - 2).$$

Taking partial derivatives, we get

$$\boldsymbol{\nabla} L = \begin{bmatrix} 6x_1 + x_2 - x_3 - 2\mu \\ x_1 - x_2 + 2x_3 + \mu \\ -x_1 + 2x_2 - x_3 - \mu \\ -(2x_1 - x_2 + x_3 - 2) \end{bmatrix}$$

Solving $\nabla L = \mathbf{0}$ is slightly tedious, but straightforward

$$\left[\begin{array}{cccc|c} 6 & 1 & -1 & -2 & 0 \\ 1 & -1 & 2 & 1 & 0 \\ -1 & 2 & -1 & -1 & 0 \\ -2 & 1 & -1 & 0 & -2 \end{array}\right]$$

$$\sim \left[\begin{array}{cccc|c} 1 & -1 & 2 & 1 & 0 \\ 6 & 1 & -1 & -2 & 0 \\ -1 & 2 & -1 & -1 & 0 \\ -2 & 1 & -1 & 0 & -2 \end{array}\right] \quad \text{Swap } R_1 \ \& \ R_2$$

$$\sim \left[\begin{array}{cccc|c} 1 & -1 & 2 & 1 & 0 \\ 0 & 7 & -13 & -8 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 3 & 2 & -2 \end{array}\right]$$

$$\sim \left[\begin{array}{cccc|c} 1 & -1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 7 & -13 & -8 & 0 \\ 0 & -1 & 3 & 2 & -2 \end{array}\right] \quad \text{Swap } R_2 \ \& \ R_3$$

$$\sim \left[\begin{array}{cccc|c} 1 & -1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & -20 & -8 & 0 \\ 0 & 0 & 4 & 2 & -2 \end{array}\right]$$

$$\sim \left[\begin{array}{cccc|c} 1 & -1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0.4 & 0 \\ 0 & 0 & 0 & 0.4 & -2 \end{array}\right]$$

Hence there is a stationary point when $\mu = -5, x_3 = 2, x_2 = -2, x_1 = -1$. Substituting this into the objective gives a value of $f = -5$.

(b) If we let x_1 be the basic variable, then $B = [2]$ and $N = \begin{bmatrix} -1 & 1 \end{bmatrix}$. Hence

$$Z = \begin{bmatrix} -B^{-1}N \\ I \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Now the Hessian of the original objective function is

$$\nabla^2 f = \begin{bmatrix} 6 & 1 & -1 \\ 1 & -1 & 2 \\ -1 & 2 & -1 \end{bmatrix}$$

Thus, the reduced Hessian is

$$
\begin{aligned}
Z^T \nabla^2 f Z &= \begin{bmatrix} \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 & 1 & -1 \\ 1 & -1 & 2 \\ -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & -4 \\ -\frac{1}{2} & \frac{3}{2} \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} \\
&= \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{3}{2} \end{bmatrix}.
\end{aligned}
$$

Checking the characteristic equation of this matrix:

$$
\begin{aligned}
\det \begin{bmatrix} \frac{3}{2} - \lambda & -\frac{1}{2} \\ -\frac{1}{2} & \frac{3}{2} - \lambda \end{bmatrix} &= \left(\frac{3}{2} - \lambda\right)^2 - \frac{1}{4} \\
&= \lambda^2 - 3\lambda + 2
\end{aligned}
$$

Setting this to zero gives $\lambda = 1$ and $\lambda = 2$ as the eigenvalues. Hence the point we found earlier is a local minimum for the problem (and actually a global minimum since $Z^T \nabla^2 f Z$ did not depend on $\mathbf{x}$ — the reduced function is strictly convex).