# Brief solutions to (some) problems in Chapter 2

Mattias Wahde

**2.3** The hessian H takes the form

$$H = \begin{pmatrix} 4 & -2 \\ -2 & 2 \end{pmatrix} \tag{1}$$

The eigenvalues are obtained from the determinant equation $\det(H - \lambda I) = 0$ that, in this case, becomes

$$(4-\lambda)(2-\lambda) - 4 = \lambda^2 - 6\lambda + 4 = 0, \tag{2}$$

with the solutions $\lambda_{1,2} = 3 \pm \sqrt{5} > 0$. Since $H$ is positive definite, $f$ is convex.

**2.5** The direction of steepest descent is given by the negative gradient $-\nabla f$. The gradient equals

$$\nabla f = (4x_1^3 + 3x_2 - 2, 3x_1 + 8x_2^3)^{\mathrm{T}}. \tag{3}$$

At the point $(2,2)^{\mathrm{T}}$, this vector equals $(36,70)^{\mathrm{T}}$. Normalizing this vector, and taking into account that the search direction is towards the *negative* gradient, the direction vector $\mathbf{r} = -(0.457, 0.889)^{\mathrm{T}}$ is obtained. Note that there is no reason to normalize the (negative) gradient for using it in the gradient descent method, but in this problem the search *direction* (i.e. a normalized vector) was specifically asked for, hence the computation of the normalized vector.

The next iterate is obtained as $\mathbf{x}_1 = \mathbf{x}_0 - \eta\nabla f(\mathbf{x}_0)$. Plotting $f(\mathbf{x}_1) \equiv f(2 - 36\eta, 2 - 70\eta) \equiv \phi(\eta)$ and applying line search (using e.g. interval halving) one finds that the minimum occurs at $\eta \approx 0.03584$. The point reached is $(0.710, -0.509)^{\mathrm{T}}$.

**2.8** With $f(x) = x^4 - x^3 + x^2 - x + 1$ one gets $f'(x) = 4x^3 - 3x^2 + 2x - 1$ and $f"(x) = 12x^2 - 6x + 2$. Inserting into the Newton-Raphson method one thus finds

$$x_{j+1} = x_j - \frac{4x_j^3 - 3x_j^2 + 2x_j - 1}{12x_j^2 - 6x_j + 2}. \tag{4}$$

Iterating from $x = x_0 = 1$, the sequence $(1, 0.75, 0.63235, \ldots, 0.6058296)$ is obtained, so that $x* = 0.6058296$. It is easy to verify that $f'(x*) \approx 0$. Also, $f"(x*) \approx 2.77 > 0$, so the stationary point is indeed a minimum. In order to study global optimality, one must investigate the Hessian, which in (note!) *one* dimension is equal to $f"(x)$, and the eigenvalue is therefore the value of $f"(x)$ itself. Here, $f"(x)$ is a parabola with a (single) stationary point (which is clearly a minimum) at the point where *its* derivative is zero, namely at $x = 1/4$. At this point $f"(x)$ takes its minimum value of $5/4$. Thus, $f"(x)$ (and, therefore the eigenvalue of the Hessian) is larger than 0 everywhere, so that the function is convex, and the local minimum is therefore also the global minimum.

**2.9** In this case, the function $L$ will be

$$L = -x_1 x_2 + \lambda(x_1^2 + 2x_2^2 - 1). \tag{5}$$

Setting the gradient to zero, one obtains the equations

$$\frac{\partial L}{\partial x_1} = -x_2 + 2\lambda x_1 = 0, \tag{6}$$

$$\frac{\partial L}{\partial x_2} = -x_1 + 4\lambda x_2 = 0, \tag{7}$$

and

$$\frac{\partial L}{\partial \lambda} = x_1^2 + 2x_2^2 - 1 = 0, \tag{8}$$

From Eqs. 6 and 7 one can note that the trivial solution $x_1 = x_2 = 0$ is not possible, as it does not fulfil the constraints. Thus, with $x_1 \neq 0$ and $x_2 \neq 0$, one finds

$$\lambda = \frac{x_2}{2x_1} = \frac{x_1}{4x_2}, \tag{9}$$

so that

$$2x_2^2 = x_1^2. \tag{10}$$

Thus, $x_1 = \pm\sqrt{2}x_2$. From the constraint equation (Eq. 8) one then finds $x_2 = \pm 1/2$, $x_1 = \pm 1/\sqrt{2}$. Checking the four possibilities, one finds that $f$ takes its minimum value of $-1/2\sqrt{2}$ at $(x_1, x_2)^{\mathrm{T}} = \pm(1/\sqrt{2}, 1/2)^{\mathrm{T}}$.