4. I denna uppgift är $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$,

där $\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}$, $\mathbf{c} = \begin{pmatrix} -1 \\ -2 \\ -4 \\ -8 \end{pmatrix}$ och $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ = variabelvektorn.

(a) Bestäm en punkt $\hat{\mathbf{x}} \in I\!R^4$ som minimerar $f(\mathbf{x})$ under bivillkoret att $x_1 + x_2 + x_3 + x_4 = 1$. .... (6p)

(b) Motivera hur du vet att det verkligen är en minpunkt du har bestämt, och inte t ex en sadelpunkt. Kända satser får användas. .... (1p)

5. Betrakta följande ickelinjära minsta-kvadratproblemet i variabelvektorn $\mathbf{x} \in I\!R^2$:

$$\text{minimera } f(\mathbf{x}) = \tfrac{1}{2}(h_1(\mathbf{x})^2 + h_2(\mathbf{x})^2 + h_3(\mathbf{x})^2 + h_4(\mathbf{x})^2), \text{ där}$$

$$\begin{aligned} h_1(\mathbf{x}) &= (x_1+2)^2 + x_2^2 - 5, \\ h_2(\mathbf{x}) &= (x_1-2)^2 + x_2^2 - 3, \\ h_3(\mathbf{x}) &= x_1^2 + (x_2-2)^2 - 6, \\ h_4(\mathbf{x}) &= x_1^2 + (x_2+2)^2 - 2. \end{aligned}$$

(a) Genomför en iteration med Gauss-Newtons metod utgående från $\mathbf{x}^{(1)} = (0,0)^\mathsf{T}$. Kontrollera att din erhållna punkt $\mathbf{x}^{(2)}$ uppfyller $f(\mathbf{x}^{(2)}) < f(\mathbf{x}^{(1)})$. .... (5p)

(b) Ett alternativ vore förstås Newtons metod för att minimera $f(\mathbf{x})$. Visa att på just detta problem, och från just denna startpunkt $\mathbf{x}^{(1)}$, så blir nästa iterationspunkt $\mathbf{x}^{(2)}$ densamma med Newtons metod som med Gauss-Newtons metod. .... (2p)

(c) Är målfunktionen $f(\mathbf{x})$ en konvex funktion på hela $I\!R^2$? Motivera svaret. (2p)

6. Antag att man vill bestämma den punkt i $I\!R^3$ som ligger närmast punkten $(0,0,2)^\mathsf{T}$ bland alla punkter i $I\!R^3$ som ligger på högst avståndet 1 från såväl punkten $(1,0,0)^\mathsf{T}$ som från punkten $(0,1,0)^\mathsf{T}$. Detta problem kan formuleras på följande sätt:

$$\begin{aligned} \text{minimera } \quad & x_1^2 + x_2^2 + (x_3-2)^2 \\ \text{då } \quad & (x_1-1)^2 + x_2^2 + x_3^2 \le 1, \\ & x_1^2 + (x_2-1)^2 + x_3^2 \le 1. \end{aligned}$$

(a) Ställ upp KKT-villkoren för problemet. .... (1p)

(b) Avgör om $\mathbf{x} = (0\,,\,0\,,\,2)^\mathsf{T}$ är en KKT-punkt. (Dvs en punkt som tillsammans med skalärer y_i uppfyller samtliga KKT-villkor.) .... (1p)

(c) Avgör om $\mathbf{x} = \left(\frac{1}{3}, \frac{1}{3}, \frac{2}{3}\right)^\mathsf{T}$ är en KKT-punkt. .... (5p)

(d) Finns det någon globalt optimal lösning till problemet? Motivera svaret utförligt. Kända satser får användas. .... (2p)

Lycka till!

Uppgift 5.(a)

Vi har att $\mathbf{h}(\mathbf{x}) = \begin{pmatrix} (x_1+2)^2 + x_2^2 - 5 \\ (x_1-2)^2 + x_2^2 - 3 \\ x_1^2 + (x_2-2)^2 - 6 \\ x_1^2 + (x_2+2)^2 - 2. \end{pmatrix}$ och $\nabla\mathbf{h}(\mathbf{x}) = \begin{bmatrix} 2x_1+4 & 2x_2 \\ 2x_1-4 & 2x_2 \\ 2x_1 & 2x_2-4 \\ 2x_1 & 2x_2+4 \end{bmatrix}$.

Vi ska starta i $\mathbf{x}^{(1)} = (0,\,0)^\mathsf{T}$. Då är $\mathbf{h}(\mathbf{x}^{(1)}) = \begin{pmatrix} -1 \\ 1 \\ -2 \\ 2 \end{pmatrix}$ och $\nabla\mathbf{h}(\mathbf{x}) = \begin{bmatrix} 4 & 0 \\ -4 & 0 \\ 0 & -4 \\ 0 & 4 \end{bmatrix}$.

I Gauss-Newtons metod ska man lösa systemet $\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\nabla\mathbf{h}(\mathbf{x}^{(1)})\mathbf{d} = -\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\mathbf{h}(\mathbf{x}^{(1)})$,

dvs $\begin{bmatrix} 32 & 0 \\ 0 & 32 \end{bmatrix}\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 8 \\ -16 \end{pmatrix}$, med lösningen $\mathbf{d}^{(1)} = \begin{pmatrix} 1/4 \\ -1/2 \end{pmatrix}$.

Vi prövar $t_1 = 1$, så att $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = \begin{pmatrix} 1/4 \\ -1/2 \end{pmatrix}$. Då blir

$$\mathbf{h}(\mathbf{x}^{(2)}) = \begin{pmatrix} (9/4)^2 + (1/2)^2 - 5 \\ (7/4)^2 + (1/2)^2 - 3 \\ (1/4)^2 + (5/2)^2 - 6 \\ (1/4)^2 + (3/2)^2 - 2. \end{pmatrix} = \begin{pmatrix} 5/16 \\ 5/16 \\ 5/16 \\ 5/16 \end{pmatrix} \quad \text{och } f(\mathbf{x}^{(2)}) = 25/128 < 5 = f(\mathbf{x}^{(1)}),$$

så steget $t_1 = 1$ gick bra. Därmed har vi utfört en iteration med Gauss-Newtons metod.

Uppgift 5.(b)

Vid Newtons metod, utgående från samma startpunkt $\mathbf{x}^{(1)} = (0,\,0)^\mathsf{T}$ som ovan, ska man i stället lösa ekvationssystemet $\mathbf{F}(\mathbf{x}^{(1)})\mathbf{d} = -\nabla f(\mathbf{x}^{(1)})^\mathsf{T}$, där $\mathbf{F}(\mathbf{x}^{(1)})$ är Hessianen av f i $\mathbf{x}^{(1)}$.

För ickelinjära minsta-kvadratproblem kan detta system skrivas

$$\left(\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\nabla\mathbf{h}(\mathbf{x}^{(1)}) + \sum_{i=1}^{m} h_i(\mathbf{x}^{(1)})\mathbf{H}_i(\mathbf{x}^{(1)})\right)\mathbf{d} = -\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\mathbf{h}(\mathbf{x}^{(1)}),$$

där $\mathbf{H}_i(\mathbf{x}^{(1)})$ är Hessianen av h_i i $\mathbf{x}^{(1)}$.

I vårt fall är $\mathbf{H}_i(\mathbf{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ för $i = 1, 2, 3, 4$, oberoende av $\mathbf{x}$.

$$\text{Då blir } \sum_{i=1}^{m} h_i(\mathbf{x}^{(1)})\mathbf{H}_i(\mathbf{x}^{(1)}) = \left(\sum_{i=1}^{m} h_i(\mathbf{x}^{(1)})\right)\mathbf{H}_i(\mathbf{x}^{(1)}) = 0\cdot\mathbf{H}_i(\mathbf{x}^{(1)}) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Därmed är Newton-systemet $\mathbf{F}(\mathbf{x}^{(1)})\mathbf{d} = -\nabla f(\mathbf{x}^{(1)})^\mathsf{T}$ identiskt med Gauss-Newton-systemet $\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\nabla\mathbf{h}(\mathbf{x}^{(1)})\mathbf{d} = -\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\mathbf{h}(\mathbf{x}^{(1)})$.
Därför blir nästa iterationspunkt $\mathbf{x}^{(2)}$ densamma för bägge metoderna.

Uppgift 5.(c) Det gäller här att avgöra om $\mathbf{F}(\mathbf{x}))$ är positivt semidefinit för alla $\mathbf{x} \in I\!R^n$. En del kalkyler visar att $\mathbf{F}(\mathbf{x})$ faktiskt är positivt definit för alla $\mathbf{x} \in I\!R^n$, vilket betyder att f inte bara är konvex utan till och med strikt konvex på hela $I\!R^n$.