Exercise Sheet 9.

Hand in due to: **Friday, 26/05/2023, 14:00**

Exercise 1 (Low-rank matrices | 4 points).

We define for $p \in \mathbb{N}$ the set of $2^p \times 2^p$ matrices of rank 1 as

$$\mathcal{R}_p = \{\mathbf{A} = \mathbf{x}\mathbf{y}^T \in \mathbb{R}^{2^p \times 2^p}\}.$$

Show that:

(a) The matrix-vector multiplication of $\mathbf{R} \in \mathcal{R}_p$ with $\mathbf{x} \in \mathbb{R}^n$ requires $N_{MV}(p) = 3n - 1$ operations.

(b) The matrix-matrix multiplication of $\mathbf{R}_1, \mathbf{R}_2 \in \mathcal{R}_p$ requires also $N_{\mathbf{R}\cdot\mathbf{R}}(p) = 3n - 1$ operations.

Exercise 2 (Hierarchical matrices | 4 points).

The set of hierarchical matrices $\mathcal{H}_k$ is recursively defined as

$$\mathcal{H}_0 := \mathbb{R}^{1\times 1},$$
$$\mathcal{H}_k := \left\{ \mathbf{H} = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{bmatrix} \in \mathbb{R}^{2^k \times 2^k} \text{ with } \mathbf{H}_{11}, \mathbf{H}_{22} \in \mathcal{H}_{k-1} \text{ and } \mathbf{H}_{12}, \mathbf{H}_{21} \in \mathcal{R}_{k-1} \right\}.$$

Information: The matrix addition of two matrices $\mathbf{R}_1, \mathbf{R}_2 \in \mathcal{R}_p$ is in general not in $\mathcal{R}_p$, because the rank may increase. Therefore, a formatted addition is used. This forms the addition $\mathbf{R}_1 + \mathbf{R}_2$ and then performs the best approximation by a matrix of rank 1. If we denote by $B_{\mathcal{R}_p}$ the best approximation in $\mathcal{R}_p$, then the formatted addition has the form

$$\mathbf{R}_1 \oplus_1 \mathbf{R}_2 := B_{\mathcal{R}_p}(\mathbf{R}_1 + \mathbf{R}_2).$$

The formatted addition in $\mathcal{R}_p$ needs $N_{\mathbf{R}+\mathbf{R}} = 18n + 29$ operations.

We define the formatted matrix addition for $\mathbf{G}, \mathbf{H} \in \mathcal{H}_k$ recursively by

$$\mathbf{G} \oplus_1 \mathbf{H} := \begin{bmatrix} \mathbf{G}_{11} \oplus_1 \mathbf{H}_{11} & \mathbf{G}_{12} \oplus_1 \mathbf{H}_{12} \\ \mathbf{G}_{21} \oplus_1 \mathbf{H}_{21} & \mathbf{G}_{22} \oplus_1 \mathbf{H}_{22} \end{bmatrix}.$$

Let $p \in \mathbb{N}$ and $n = 2^p$. Show:

(a) The matrix-vector multiplication of $\mathbf{H} \in \mathcal{H}_p$ with $\mathbf{x} \in \mathbb{R}^n$ requires $N_{MV}(p) = 4n \log_2 n - n + 2$ operations.

(b) The formatted addition of $\mathbf{G}, \mathbf{H} \in \mathcal{H}_p$ and that of $\mathbf{H} \in \mathcal{H}_p$ with $\mathbf{R} \in \mathcal{R}_p$ require $N_{\mathbf{H}+\mathbf{H}}(p) = N_{\mathbf{H}+\mathbf{R}}(p) = 18n \log_2 n + 59n - 58$ operations.

(c) The matrix-matrix multiplication of $\mathbf{G}, \mathbf{H} \in \mathcal{H}_p$ requires $N_{\mathbf{H}\cdot\mathbf{H}}(p) = 13n \log_2^2 n + 65n \log_2 n - 51n + 52$ operations. For the occurring additions, formatted additions shall be used here.

Exercise 3 (Constrained optimization | 4 points).

For $\mathbf{x} \in \mathbb{R}^3$ consider the constrained optimization problem

$$\min_{\mathbf{x}\in\mathbb{R}^3} \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x} \quad \text{subject to} \quad \mathbf{C}^\top \mathbf{x} = \boldsymbol{\gamma}, \quad \mathbf{D}^\top \mathbf{x} \geq \boldsymbol{\delta},$$

and consider the Lagrangian

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\lambda}) := \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x} - \boldsymbol{\mu}^\top(\mathbf{C}^\top \mathbf{x} - \boldsymbol{\gamma}) - \boldsymbol{\lambda}^\top(\mathbf{D}^\top \mathbf{x} - \boldsymbol{\delta}).$$

Show that if

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \qquad \mathbf{D} = \mathbf{I}, \quad \boldsymbol{\gamma} = \boldsymbol{\delta} = \mathbf{0},$$

then the triple $(\mathbf{x}^\star, \mu^\star, \boldsymbol{\lambda}^\star)$ given by

$$\mathbf{x}^\star = \frac{1}{9}(1, 1, 0)^\top, \quad \mu^\star = -\frac{5}{9}, \quad \boldsymbol{\lambda}^\star = \frac{1}{9}(0, 0, 1)^\top,$$

satisfies the Karush-Kuhn-Tucker (KKT) conditions

$$\nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x}^\star, \boldsymbol{\mu}^\star, \boldsymbol{\lambda}^\star) = \mathbf{0}, \quad \mathbf{C}^\top \mathbf{x}^\star = \boldsymbol{\gamma}, \quad \mathbf{D}^\top \mathbf{x}^\star \geq \boldsymbol{\delta}, \quad \boldsymbol{\lambda} \geq \mathbf{0}, \quad \boldsymbol{\lambda}^\top(\mathbf{D}^\top \mathbf{x}^\star - \boldsymbol{\delta}) = \mathbf{0}.$$

Exercise 4 (Dual problem | 4 points).

Suppose that f and $-c_i$, $i = 1, 2, \dots, m$, are convex and continuously differentiable functions on $\mathbb{R}^n$. For $\mathbf{x} \in \mathbb{R}^n$ consider the constrained optimization problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \quad \text{subject to} \quad \mathbf{c}(\mathbf{x}) \geq \mathbf{0}, \tag{1}$$

where $\mathbf{c}(\mathbf{x}) := \big(c_1(\mathbf{x}), c_2(\mathbf{x}), \dots, c_m(\mathbf{x})\big)^\top$. The Lagrangian with Lagrange multiplier $\boldsymbol{\lambda} \in \mathbb{R}^m$ is

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda}^\top \mathbf{c}(\mathbf{x}).$$

We define the dual objective function $q : \mathbb{R}^m \to \mathbb{R}$ as follows

$$q(\boldsymbol{\lambda}) := \inf_{\mathbf{x}\in\mathbb{R}^m} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}).$$

Then, the dual problem is defined as follows

$$\max_{\lambda\in\mathbb{R}^m} q(\boldsymbol{\lambda}) \quad \text{subject to} \quad \boldsymbol{\lambda} \geq \mathbf{0}. \tag{2}$$

(a) Suppose that $\mathbf{x}^*$ is a solution of (1). Show that any $\boldsymbol{\lambda}^*$, for which $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ satisfies the KKT conditions, is a solution of (2).

(b) Suppose that $\mathbf{x}^*$ is a solution of (1) and $\boldsymbol{\lambda}^*$ satisfies the KKT conditions. Suppose that $\tilde{\boldsymbol{\lambda}}$ solves (2) and $\tilde{\mathbf{x}} = \operatorname{arginf}_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \tilde{\boldsymbol{\lambda}})$. Assume further that $\mathcal{L}(\cdot, \tilde{\boldsymbol{\lambda}})$ is a strictly convex function. Show that $\mathbf{x}^* = \tilde{\mathbf{x}}$, i.e., $\tilde{\mathbf{x}}$ is the unique solution of (1), and $f(\mathbf{x}^*) = \mathcal{L}(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}})$.