# Minimization in $\mathbb{R}^N$ : Gradient Methods

Let us consider a functional $\mathcal{J}(\boldsymbol{X}) : \mathbb{R}^N \mapsto \mathbb{R}$, defined as $\mathcal{J}(\boldsymbol{X}) = \frac{1}{2}\boldsymbol{X} \cdot \left[\underline{\mathcal{C}}\boldsymbol{X}\right] + \boldsymbol{B}_0 \cdot \boldsymbol{X}$, where $\underline{\mathcal{C}}$ is a $N \times N$ matrix and $\boldsymbol{B}_0$ is a vector of size $N$. For example

$$\underline{\mathcal{C}} = -(N+1)\begin{pmatrix} -2 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & -2 & 1 \\ 0 & \cdots & 0 & 1 & -2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{B}_0 = -\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ (N+1) \end{pmatrix}$$

and in this case the minimum of $\mathcal{J}$ is $\boldsymbol{U} \in \mathbb{R}^N$ with $\boldsymbol{U}_j = \frac{j}{N+1}$ In order to solve the minimization problem in $\mathbb{R}^N$ it is defined, step by step from a gigen vector $\boldsymbol{V}^0$ (for example $\boldsymbol{V}^0 \equiv 0$), positions $\boldsymbol{V}^k \in \mathbb{R}^N$ such as $\mathcal{J}(\boldsymbol{V}^{k+1}) \leq \mathcal{J}(\boldsymbol{V}^k)$
For the gradient methods

$$\boldsymbol{V}^{k+1} = \boldsymbol{V}^k + \alpha_k \boldsymbol{d}^k \quad \text{where} \quad \boldsymbol{d}^k = -\nabla\mathcal{J}(\boldsymbol{V}^k).$$

In the curent case, the gradient can be computed either analiticaly or numerically.

**Gradient method with Constant step $\alpha_k = \alpha$.** In a fortran program, compute $\boldsymbol{V}_\alpha^K$, for $N = 200$, $K = N/2$, using values $\alpha = 10, 1, 10^{-1}, 10^{-2}, ...$, For the defined example, plot $\|\boldsymbol{U} - \boldsymbol{V}_\alpha^K\|$ as a function of $\alpha$ and define, in this case, an approximated convergence criteria.

**Gradient method with optimal step.** For a given vector $\boldsymbol{V}_k$ and a descent direction $\boldsymbol{d}_k \in \mathbb{R}^N$, we define a functional $f_k(\rho) : \mathbb{R} \mapsto \mathbb{R}$ as $f_k(\rho) = \mathcal{J}(\boldsymbol{V}_k + \rho\boldsymbol{d}_k)$ Then for the optimal step, $\alpha_k$ is the minimum of $f_k(\rho)$ on an interval $[0, b]$ with $0 < b$. In a fortran program, compute $\boldsymbol{V}_b^K$, for $N = 200$, $K = N/2$, using values $b = 10^{-2}, 10^{-1}, 1, 10^2, 10^3, ...$, For the defined example, plot $\|\boldsymbol{U} - \boldsymbol{V}_b^K\|$ as a function of $b$.

**Gradient method with optimal step and equality constraint.** In this item we consider the minimization in a sub-space $\mathcal{E}$ defined by vector $\boldsymbol{O}_e$ the normal $\boldsymbol{n}$ with $\|\boldsymbol{n}\| \neq 0 : \mathcal{X} \in \mathcal{E} \iff (\mathcal{X} - \boldsymbol{O}_e) \cdot \boldsymbol{n} = 0$. Propose a modification of the optimal step method to approximate the minimum of $J$ on $\mathcal{E}$.

**Gradient method with optimal step and inequality constraint.** In this item we consider the minimization in a half space $\mathcal{E}$ defined by vector $\boldsymbol{O}_e$ the normal $\boldsymbol{n}$ with $\|\boldsymbol{n}\| \neq 0 : \mathcal{X} \in \mathcal{E} \iff (\mathcal{X} - \boldsymbol{O}_e) \cdot \boldsymbol{n} \geq 0$. Propose a modification of the optimal step method to approximate the minimum of $J$ on $\mathcal{E}$.

**Gradient method with optimal step and strict inequality constraint.** What about the case of $\mathcal{E}$ defined by vector $\boldsymbol{O}_e$ the normal $\boldsymbol{n}$ with $\|\boldsymbol{n}\| \neq 0 : \mathcal{X} \in \mathcal{E} \iff (\mathcal{X} - \boldsymbol{O}_e) \cdot \boldsymbol{n} > 0$ ?

*for the minimization of $f_k(\rho)$ use the* `Dichotomous` *or the* `GoldSection` *algorithm.*