ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

*Prof. D. Kressner*
*M. Steinlechner*

## 1 ► Rosenbrock function with equality constraint

In this exercise, we want to find the minimum of the Rosenbrock function subject to an equality constraint:

$$\min_{\mathbf{x}\in\mathbb{R}^2} \quad f(\mathbf{x}) := 100\big(x_2 - x_1^2\big)^2 + (1-x_1)^2$$
$$\text{subject to:} \quad x_1^2 + x_2^2 - 2 = 0.$$

(a) Write down the KKT conditions for this system, equivalent to finding the zeros of a nonlinear function. Apply the Newton method to derive equation (5.34) in the script for this problem. Finally, substitute $\mu^{k+1}$ to arrive at the linear system (5.36) corresponding to the local quadratic program.

(b) Find the solution of the constraint optimization problem by a simple local SQP algorithm: Beginning from the initial guess $(x_1^0, x_2^0, \mu^0) = (2.5, 5, 1)$, solve the quadratic program around the current iterate $(x_1^k, x_2^k, \mu^k)$ and update: $\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{p}^k$. Compare to the initial guess $(x_1^0, x_2^0, \mu^0) = (0.75, 5, 1)$. What do you observe?

## 2 ► Subdifferential

In this exercise, we want to explicitely calculate the subdifferential, the set of all subgradients, of simple functions.

a) $f(x) = |x|$ at $x_0 = 1$ and $x_0 = 0$.

b) $f(x) = \max\left\{0, \frac{1}{2}(x^2-1)\right\}$ at $x_0 = \pm 1$.

c) $f(x) = -\sqrt{x}$ with $x \in \{y \in \mathbb{R} \,|\, y \geq 0\}$ at $x_0 = 0$.

*Hint: Visualize the corresponding subgradients at $x_0$ graphically – consider the possible "tangents" to the function at $x_0$!*

## 3 ► Proximal operators

a) Consider the $\ell_1$-norm of a vector, $g(x) = \|x\|_1 = \sum_{i=1}^n |x_i|$ for $x \in \mathbb{R}^n$ and let $\lambda > 0$ be a parameter. Show that we can compute the proximal operator for this function explicitly in the following form (for the $i$th component of $\text{prox}_{\lambda g}(x)$):

$$\Big(\text{prox}_{\lambda g}(x)\Big)_i = \begin{cases} x_i - \lambda & \text{if } x_i > \lambda, \\ x_i + \lambda & \text{if } x_i < -\lambda, \\ 0 & \text{if } -\lambda \leq x_i \leq \lambda. \end{cases},$$

which can be written in short form as $\Big(\text{prox}_{\lambda g}(x)\Big)_i = \max\{|x_i| - \lambda, 0\}\,\text{sign}(x_i)$.

This operator is also known as the *shrinkage* (soft thresholding) operator

$$S_\lambda(x) := \text{prox}_{\lambda g}(x).$$

*b) Let $A$ be an $n \times n$ matrix and $g(A) := \sigma_1(A) + \cdots + \sigma_n(A)$, where $\sigma_j$ is the $j$th singular value of A. What is $\text{prox}_{\lambda g}(A)$ ?

**Note:** In the definition of $\text{prox}_{\lambda g}(A)$, the 2-norm of the vector needs to be replaced by the Frobenius norm of the matrix.