Course of Numerical Methods for Engineering
Lab 10

Luca Bonaventura, Tommaso Benacchio

MOX - Politecnico di Milano

PHYS-ENG, A.Y. 2020-21

16/11/2020

Topic of this session:

- **Methods for function minimization**

Definition

*(***The modified gradient method***) Let* $x^{(0)}$ *an initial approximation for the minimum of* ϕ *and let* γ *an arbitrary positive real number. The modified gradient or modified steepest descent method is defined by*

$$x^{(k+1)} = x^{(k)} - \gamma_k \nabla\phi(x^{(k)}) \qquad k \geq 0,$$

where $\gamma_k = \gamma/2^m$ *and* m *is the smallest integer such that*

$$\phi\left(x^{(k)} - \frac{\gamma}{2^m}\nabla\phi(x^{(k)})\right) - \phi(x^{(k)}) \leq -\frac{1}{2}\frac{\gamma}{2^m}\|\nabla\phi(x^{(k)}))\|_2^2.$$

Definition

*(**The Newton method for function minimization**) Let $\phi \in \mathcal{C}^2(\Omega)$ and let*

$$\mathrm{f}(\mathrm{x}) = \left[\frac{\partial \phi}{\partial x_1}(\mathrm{x}), \ldots, \frac{\partial \phi}{\partial x_n}(\mathrm{x})\right]^T = \nabla \phi(\mathrm{x}).$$

Let $\mathrm{x}^{(0)}$ an initial approximation for the minimum of ϕ. The Newton method for the minimization of ϕ is defined as

$$\begin{aligned} \mathrm{H}(\mathrm{x}^{(k)})\boldsymbol{\delta}\mathrm{x}^{(k)} &= -\mathrm{f}(\mathrm{x}^{(k)}) = -\nabla\phi(\mathrm{x}^{(k)}) \\ \mathrm{x}^{(k+1)} &= \mathrm{x}^{(k)} + \boldsymbol{\delta}\mathrm{x}^{(k)} \quad k \geq 0. \end{aligned}$$

Definition

*(***The BFGS quasi-Newton method for function minimization***) Let $\phi \in \mathcal{C}^2(\Omega)$ and let $\mathsf{f}(\mathsf{x}) = \nabla\phi(\mathsf{x})$. Let $\mathsf{x}^{(0)}$ an initial approximation for the minimum of ϕ and let H_0 denote a symmetric and positive definite matrix that is an initial approximation of the Hessian of ϕ at $\mathsf{x}^{(0)}$. The BFGS quasi-Newton method for the minimization of ϕ is then defined as*

$$
\begin{aligned}
\mathsf{H}_k \boldsymbol{\delta}\mathsf{x}^{(k)} &= -\mathsf{f}(\mathsf{x}^{(k)}) = -\nabla\phi(\mathsf{x}^{(k)}) \\
\mathsf{x}^{(k+1)} &= \mathsf{x}^{(k)} + \gamma_k \boldsymbol{\delta}\mathsf{x}^{(k)} = \mathsf{x}^{(k)} + \boldsymbol{\delta}_k \\
& \quad \gamma_k \text{ is such that } \phi(\mathsf{x}^{(k)} + \gamma_k \boldsymbol{\delta}\mathsf{x}^{(k)}) = \min_{\gamma > 0} \phi(\mathsf{x}^{(k)} + \gamma \boldsymbol{\delta}\mathsf{x}^{(k)}) \\
\boldsymbol{\nu}_k &= \mathsf{f}(\mathsf{x}^{(k+1)}) - \mathsf{f}(\mathsf{x}^{(k)}) = \nabla\phi(\mathsf{x}^{(k+1)}) - \nabla\phi(\mathsf{x}^{(k)}) \\
\mathsf{H}_{k+1} &= \mathsf{H}_k + \frac{\boldsymbol{\nu}_k \boldsymbol{\nu}_k^T}{\boldsymbol{\nu}_k^T \boldsymbol{\delta}_k} - \frac{\mathsf{H}_k \boldsymbol{\delta}_k \boldsymbol{\delta}_k^T \mathsf{H}_k}{\boldsymbol{\delta}_k^T \mathsf{H}_k \boldsymbol{\delta}_k}.
\end{aligned}
$$

Finding γ

The minimization of $\phi(\mathrm{x}^{(k)} + \gamma\boldsymbol{\delta}\mathrm{x}^{(k)})$ can be carried out at each step in an approximate fashion, for example by taking $\gamma_k^{(0)} = 0$, checking for for some increasing sequence $\gamma_k^{(l)},\ l \geq 0$, that
$\phi(\mathrm{x}^{(k)} + \gamma_k^{(l+1)}\boldsymbol{\delta}\mathrm{x}^{(k)}) < \phi(\mathrm{x}^{(k)} + \gamma_k^{(l)}\boldsymbol{\delta}\mathrm{x}^{(k)})$ and stopping as soon as the previous inequality is violated, setting $\gamma_k = \gamma_k^{(l)}$.

Exercise 1

Consider the Rosenbrock banana function $\phi(x,y) = (a-x)^2 + b(y-x^2)^2$ **with** $a=1, b=4$ **on the domain** $[0,4]\times[0,4]$.
Find its global minimum minimum $(1,1)$ **using**

(a) **the MATLAB command** `fminsearch`**;**

(b) **the modified gradient method;**

(c) **the Newton method for function minimization;**

(d) **the BFGS quasi-Newton method.**

Use the initial guesses $(0,0)$, $(2,2)$, $(1,3)$, $(3,1)$ **and a** 10^{-8} **tolerance on the function value. For the modified gradient method, start from an initial** $\gamma = 0.1$**. For the BFGS method, check for** γ **values** 0.01 **apart. Compare the number of iterations required by each method with each initial guess. Repeat the exercise for** $b = 100$**.**

Exercise 2

Consider the function $\phi(x, y) = (x^2 + y^2)^2 - 15(4x^2 + (\frac{y}{2})^2) + 20x + 5y$ **on the domain** $[-7, 7] \times [-7, 7]$. **Identify graphically appropriate initial guesses for minimization methods (use the commands** `meshgrid` **,** `mesh` **and** `contourf`**). Find its local and global minima using**

(a) **the MATLAB command** `fminsearch`**;**

(b) **the modified gradient method;**

(c) **the Newton method for function minimization;**

(d) **the BFGS quasi-Newton method.**

Compare the number of iterations required by each method with each initial guess. Use the tolerances 10^{-8} **and** 10^{-12} **(a) on the solution; and (b) on the function value.**

Exercise 3

Let A **be the** $n \times n$ **tridiagonal matrix with elements equal to 5 on the main diagonal and -1 on the first upper and lower diagonals and consider the nonlinear system:**

$$\mathrm{f}(\mathrm{x}) = \mathrm{A}\mathrm{x} + \mathrm{g}(\mathrm{x}) = 0,$$

where $\mathrm{g}(\mathrm{x}) = [-2\sin(x_1) + \pi, \ldots, -2\sin(x_n) + \pi]$.
In the case $n = 10$, **reformulate the system as the problem of minimizing** $\phi(x) = \|\mathrm{f}(\mathrm{x})\|_2^2$ **and solve it up to a** 10^{-8} **tolerance using**

(a) **the MATLAB command** `fminsearch` **;**

(b) **the modified gradient method;**

(c) **the Newton method;**

(d) **the BFGS quasi-Newton method.**

Use the initial guess $\mathrm{x}_0 = [1, \ldots, 1]^T$ **and compare the results with those obtained in the corresponding exercise of the previous exercise session. Repeat for** $n = 100$.

Hint for exercise 3: **If $\phi(\mathbf{x}) = \|\mathbf{x}\|_2^2 = \sum_k f_k^2$, then**

$$\frac{\partial \phi}{\partial x_j} = 2\sum_k \frac{\partial f_k}{\partial x_j} f_k,$$

so that $\nabla\phi(\mathbf{x}) = 2\mathbf{J}^T(\mathbf{x})\mathrm{f}(\mathbf{x})$. One then has

$$\frac{\partial^2 \phi}{\partial x_i \partial x_j} = 2\sum_k \frac{\partial}{\partial x_j}\left(\frac{\partial f_k}{\partial x_j} f_k\right) = 2\sum_k \left(\frac{\partial f_k}{\partial x_i}\frac{\partial f_k}{\partial x_j} + f_k \frac{\partial^2 f_k}{\partial x_i \partial x_j}\right)$$

so that the Hessian is:

$$\mathrm{H}_\phi = 2\mathbf{J}^T\mathbf{J} + 2\sum_k f_k \mathrm{H}_{f_k}$$