Fall 2017: Numerical Methods I
Assignment 4 (due Nov 2, 2017)

1. **[Convexity, 1+1+1pt]** A set $\mathcal{S}$ is convex, if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{S}$, also the convex combinations $\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}$ are in $\mathcal{S}$, for all $\lambda \in (0,1)$. Convex functions are usually defined over convex sets (or over all of $\mathbb{R}^n$, which is, of course, a convex set).

   (a) Show that for a convex function f, the set of global minimizers of f is a convex set.

   (b) Let $\|\cdot\|$ be any norm on $\mathbb{R}^n$. Show that $\boldsymbol{x} \mapsto \|\boldsymbol{x}\|$ is a convex function.

   (c) Let $Q \in \mathbb{R}^{n\times n}$ be symmetric and positive semidefinite, and $\boldsymbol{b} \in \mathbb{R}^n$. Prove that the function $\boldsymbol{x} \mapsto \frac{1}{2}\boldsymbol{x}^T Q\boldsymbol{x} - \boldsymbol{b}^T\boldsymbol{x}$ is convex.

2. **[Necessary and sufficient optimality conditions, 1+1+1pt]** Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x,y) = (x-y^2)(3x-y^2)$.

   (a) Show that f satisfies the first and second-order necessary optimality conditions at $\boldsymbol{x}^* = (0,0)$, but every neighborhood of $\boldsymbol{x}^*$ contains a point $\boldsymbol{z}$ with $f(\boldsymbol{z}) < f(\boldsymbol{x}^*)$.

   (b) Nevertheless, show that for every direction $\boldsymbol{d} \in \mathbb{R}^2$, $\alpha^* = 0$ is a minimizer for the restriction to the one-dimensional function $\alpha \mapsto g(\alpha) := f(\boldsymbol{x}^* + \alpha\boldsymbol{d})$.

   (c) Give an example of a twice continuously differentiable function that has a minimum at a point $\boldsymbol{x}^*$ that satisfies first and second-order necessary conditions, but does not satisfy the second-order sufficient conditions.

3. **[Modified metric in steepest descent, 2pt]** Consider $f : \mathbb{R}^n \to \mathbb{R}$ continuously differentiable, and $\boldsymbol{x} \in \mathbb{R}^n$ with $\nabla f(\boldsymbol{x}) \neq 0$. For a symmetric positive definite matrix $A \in \mathbb{R}^{n\times n}$, we define the A-weighted norm $\|\boldsymbol{y}\|_A = \sqrt{\boldsymbol{y}^T A\boldsymbol{y}}$. Derive the unit norm steepest descent direction of f in $\boldsymbol{x}$ with respect to the $\|\cdot\|_A$-norm, i.e., find the solution to the problem[1]

$$\min_{\|\boldsymbol{d}\|_A = 1} \nabla f(\boldsymbol{x})^T \boldsymbol{d}.$$

   Hint: Use the factorization $A = B^T B$ and the Cauchy-Schwarz inequality.

4. **[Gaussian mixtures interpolation/data fitting (continuation of problem 5 from the previous assignment), 1+1+2+1pt+2pt extra credit]** Let us try an approach for fitting data that avoids having to solve a *nonlinear* least squares problem. We use the same measurements as for the previous assignment, but choose a larger N and make several fixed choices for τ_j, σ_j, and only use a numerical method to find appropriate weights ω_j. We use two different values for σ_j, namely 1 and $1/2$, and 11 values for τ_j, namely $\tau_j \in \{-5, -4, \ldots, 4, 5\}$. This gives the model function

$$\varphi_2(t;\boldsymbol{x}) = \sum_{j=1}^{11} \omega_j \exp(-\frac{1}{2}(t-(j-6))^2) + \sum_{j=1}^{11} \omega_{j+11} \exp(-2(t-(j-6))^2). \quad (1)$$

[1] If f is twice differentiable with positive definite Hessian matrix, one can choose $A = \nabla^2(\boldsymbol{x})$. This shows that the Newton descent direction is the steepest descent direction in the metric where norms are weighted by the Hessian matrix.

Now, the only model parameters are $\boldsymbol{x} = (\omega_1 \ldots, \omega_{22})^T$, and since these enter linearly in φ_2, this is now a *linear* least squares problem, namely

$$\min_{\boldsymbol{x}\in\mathbb{R}^{22}} \frac{1}{2}\sum_{i=1}^{10}(\varphi_2(t_i;\boldsymbol{x}) - f_i)^2. \tag{2}$$

(a) Specify the entries of the matrix $A \in \mathbb{R}^{10\times 22}$ and the vector $\boldsymbol{f} \in \mathbb{R}^{10}$ such that (2) can be written as

$$\min_{\boldsymbol{x}\in\mathbb{R}^{22}} \frac{1}{2}\|A\boldsymbol{x} - \boldsymbol{f}\|^2.$$

(b) Comparing the number of unknowns and linear equations, we see that this is an *underdetermined* system, which does not have a unique solution $\boldsymbol{x}$. Find a solution to this underdetermined system using the built-in function (backslash in MATLAB). Amongst all of the possible solutions to this system, the built-in function should return a vector $\boldsymbol{x}$ that has some sparsity, i.e., a vector where several entries are zero, which means that many of the coefficients ω_j in (1) are zero. Because the terms with zero weight can, of course, be neglected in (1), we have found a model with small(ish) N. Let's aim at improving the result by looking at alternative approaches.

(c) To make the solution $\boldsymbol{x}$ unique, we can also consider the following modified least squares problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^{22}} \frac{1}{2}\|A\boldsymbol{x} - \boldsymbol{f}\|^2 + \frac{\gamma}{2}\|\boldsymbol{x}\|^2, \tag{3}$$

where γ is a (small) positive number. Show that the minimizer of (3) satisfies the modified normal equation

$$(A^TA + \gamma I)\boldsymbol{x} = A^T\boldsymbol{f}, \tag{4}$$

where $I \in \mathbb{R}^{22\times 22}$ denotes the identity.[2] Prove that solutions of (4) are unique for $\gamma > 0$.

(d) Solve (4) for different $\gamma > 0$ and plot the resulting model curves on top of the data points (t_i, f_i). We are told that N should be small in the model function, which would mean that many of the components in $\boldsymbol{x}$ are zero or close to zero. Is this the case for the solutions of (4) that you found?[3]

(e) Let us try to obtain a vector $\boldsymbol{x}$ that is sparser, i.e., has more zeros (and thus the model $\varphi(t)$ has a small N as desired). We add a different term to the minimization problem compared to (3), namely:

$$\min_{\boldsymbol{x}\in\mathbb{R}^{22}} \frac{1}{2}\|A\boldsymbol{x} - \boldsymbol{f}\|^2 + \gamma\|\boldsymbol{x}\|_1, \tag{5}$$

where $\|\boldsymbol{x}\|_1 = \sum_{j=1}^{22}|x_j|$ is the 1-norm. This choice tends to "favor" sparse vectors, i.e., by enforcing that $\|\boldsymbol{x}\|_1$ be small, many entries of the computed $\boldsymbol{x}$ are zero. We

[2] Hint: At the minimizer of (3), the gradient with respect to $\boldsymbol{x}$ of $1/2\|A\boldsymbol{x} - \boldsymbol{f}\|^2 + \gamma/2\|\boldsymbol{x}\|^2$ must be zero. Alternatively, you can consider the an extended least squares problem with the autmented matrix $\bar{A} = [A, \sqrt{\gamma}I]^T \in \mathbb{R}^{32\times 22}$ and use the results from class for overdetermined systems.

[3] It shouldn't be. By adding the additional term in (3), we say that we prefer vectors $\boldsymbol{x}$ with small (squared) 2-norm, but not that we are looking for solutions $\boldsymbol{x}$ that contain a large number of zeros.

would like to take derivatives of the function in (5) but, as we have seen before, this is problematic since the 1-norm is not differentiable. We thus replace (5) by

$$\min_{\boldsymbol{x}\in\mathbb{R}^{22}} \frac{1}{2}\|A\boldsymbol{x}-\boldsymbol{f}\|^2 + \gamma \sum_{j=1}^{22} \sqrt{x_j^2+\varepsilon}\,, \tag{6}$$

where ε is a (small) positive number. Show[4] that the solution of (6) satisfies the nonlinear equation

$$A^TA\boldsymbol{x} + \gamma D(\boldsymbol{x})\boldsymbol{x} = A^T\boldsymbol{f}, \tag{7}$$

where $D(\boldsymbol{x}) = \text{diag}(1/\sqrt{x_1^2+\varepsilon}, \ldots, 1/\sqrt{x_{22}^2+\varepsilon}) \in \mathbb{R}^{22\times 22}$. To solve this nonlinear equation, choose an initial guess $\boldsymbol{x}^0$ and use the fixed point-like iteration

$$(A^TA + \gamma D(\boldsymbol{x}^{k-1}))\boldsymbol{x}^k = A^T\boldsymbol{f} \quad \text{for } k = 1, 2, \ldots$$

Try different values for γ and ε[5], and plot the model functions which result from the computed parameters. Are the vectors $\boldsymbol{x}$ obtained with this method sparser? Since we had to introduce the parameter ε to make the problem differentiable, most entries in $\boldsymbol{x}$ will not exactly be zero, but they should be rather small. Based on your results, what do you think are the parameters of the Gaussian mixture used to generate the original data?

5. **[Behavior of descent methods, 6pt]** Consider the unconstrained optimization problem

$$\min f(x,y) \equiv -\cos x \cos(y/5).$$

(a) Find and classify all stationary points in the region $-\pi/2 \le x \le \pi/2, -5\pi/2 \le y \le 5\pi/2$.

(b) There is a portion of the problem region within which the Hessian matrix of $f(x,y)$ is positive definite. Give expressions for this portion. You should be able to do this analytically.

(c) Derive expressions for the search directions associated with the steepest descent and Newton methods.

(d) Write a program that performs both iterations, both without a line search and with an exact line search. Note that you will not be able to find the value of the optimal step length analytically; instead, determine it numerically.[6]

(e) Run your program for various initial guesses within the region. Verify the following:[7]

[4] Again, you can show that by computing the gradient of the function in (6) and using that it has to be zero at the minimizer.

[5] To get good results, these should be small, e.g., $\gamma \sim 10^{-3}$ and $\varepsilon \sim 10^{-9}$

[6] You can use a built-in one-dimensional minimization function (`fzero` in MATLAB). While we use exact line search here, this is usually too costly as it requires a large number of function evaluations. As we've discussed, one thus uses an inexact step size that satisfies, for instance, the Wolfe conditions to guarantee convergence to a stationary point.

[7] One way to illustrate this (non)convergence is to randomly choose initializations and draw them as dots with different colors depending on whether the method, started from that initialization, converged.

- Steepest descent converges to the minimum x^* for any starting point within the region.
- Newton's method with line search converges to the minimum only for initial points for which the Hessian matrix is positive definite.
- Newton's method without line search has an even more restricted radius of convergence.

(f) What do you observe about the convergence rate in these cases?

6. **[Globalization of Newton descent, 3+1pt]** As we have seen, the Newton direction for solving a minimization problem is only a descent direction if the Hessian matrix is positive definite. This is not always the case, in particular far from the minimizer. To guarantee a descent direction in Newton's method, a simple idea is as follows (where we choose $0 < \alpha_1 < 1$ and $\alpha_2 > 0$):

   - Compute a direction $\boldsymbol{d}^k$ by solving the Newton equation

     $$\nabla^2 f(\boldsymbol{x}^k)\boldsymbol{d}^k = -\nabla f(\boldsymbol{x}^k).$$

     If that is possible and $\boldsymbol{d}^k$ satisfies[8]

     $$\frac{-\nabla f(\boldsymbol{x}^k)^T \boldsymbol{d}^k}{\|\nabla f(\boldsymbol{x}^k)\|\|\boldsymbol{d}^k\|} \geq \min\left(\alpha_1, \alpha_2 \|\nabla f(\boldsymbol{x}^k)\|\right), \tag{8}$$

     then use $\boldsymbol{d}^k$ as descent direction.
   - Otherwise, use the steepest descent direction $\boldsymbol{d}^k = -\nabla f(\boldsymbol{x}^k)$.

   To illustrate this globalization, let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

   $$f(\boldsymbol{x}) = \frac{1}{2}(x_1^2 + x_2^2)\exp(x_1^2 - x_2^2).$$

   (a) Using the initial iterate $\boldsymbol{x}^0 = (1, 1)^T$, find a local minimum of f using the modified Newton method described above, combined with Armijo line search with backtracking.[9] Hand in a listing of your implementation.[10]

   (b) Carry out the computation also with the modified Newton matrix $\nabla^2 f(\boldsymbol{x}) + 3\boldsymbol{I}$, which has been made to be positive definite by adding a multiple of the identity. Discuss your findings.

[8] This is a condition on the angle between the negative gradient and the Newton directions, which must less than 90°. However, using this condition, the angle may approach 90° at the same speed as $\|\nabla f(\boldsymbol{x}^k)\|$ approaches zero. Recall from class that what is required to guarantee convergence for a descent method with Wolfe line search is that an infinite sum that involves the square of the right hand side in (8) and the norm of the gradient is finite.

[9] For the Armijo linesearch, one usually uses c to be rather small, e.g., $c = 10^{-4}$. Use $\alpha = 0.1$ and $\alpha_2 = 0.5$ for the constants in (8).

[10] It is sufficient to hand in a listing of the important parts of your implementation, i.e., Armijo line search and computation of the descent direction.