Math 104C Homework #5

Due on May 8 by 11:59 PM

Topics: (04/30 – 05/07) Nonlinear Least square

Video: How to submit homework on Gradescope or copy and paste (`https://youtu.be/quBWbQ5opT0`)

Update (2024-05-04 07:49 PM)

- A serious typo is corrected in Problem 2.

  $\cancel{y = c_3 + c_1 t_2} \longrightarrow y = c_3 + c_1 t^{c_2}$.

- Problem 6 has been moved to HW6 since there will be only a short time for that after covering GMRES.

I. For presentation

1. (Exploration; Analysis) Prove the following multidimensional product rule: Let $u(x_1, \ldots, x_n)$ and $v(x_1, \ldots, x_n)$ be $\mathbb{R}^n$-vector-valued functions, and let $A(x_1, \ldots, x_n)$ be an $n \times n$ matrix function. The dot product $u^T v$ is a scalar function. Then, we have, (a)

$$\nabla\left(u^T v\right) = v^T Du + u^T Dv, \quad \text{(dot product rule)}$$

and (b)

$$D(Av) = A \cdot Dv + \sum_{i=1}^{n} v_i Da_i, \quad \text{(matrix-vector product rule)}$$

where a_i denotes the i th column of A.

Here, we follow the same convention as our textbook.

- Vectors are column vectors by default.
- The gradient of a scalar function is viewed as a row vector. If $x \in R^n$, and $f : R^n \to R$, $\nabla f(x) = [\partial f/\partial x_i]_{i=1}^n$.
- D is a Jacobian operator.
  - If $x \in R^n$, and $v : R^n \to R^n$, $Dv(x)$ is a matrix whose (i, j)-component is given by $\partial v_i/\partial x_j$.
  - Or equivalently, $Dv(x)$ is a horizontal stack of gradients of each component: $\begin{bmatrix} \nabla v_1 \\ \vdots \\ \nabla v_n \end{bmatrix}$

Hint: For both (a) and (b), carefully keep track of dimensions. For (a), calculate $\partial/\partial x_i$ of $u^T v$ and organize the result in the form "(row vector) = (row vector) + (row vector)." For (b), examine (i, j)-entry of $D(Av)$ and organize the result in the form "matrix = matrix + matrix."

2. (Exploration; Core and Analysis) Find the matrix Dr needed for the application of Gauss-Newton iteration to the model-fitting problem with three data points $(t_1, y_1), (t_2, y_2), (t_3, y_3)$ translated power law $y = c_3 + c_1 t^{c_2}$.

3. (Exploration; Computation) Use the Gauss-Newton to fit the blood concentration model to the data of from the previous HW without linearization. (Additional question only for presentation) Do the same with and Levenberg-Marquardt Method with $\lambda = 1$. How are their performance different?

II. Not for presentation

The following problems are for your own study, but not for presentation. They have been already discussed during lectures.

4. (Formation) Answer the following. And give a **brief explanation** for each subproblem. (Only one or two sentences; a long explanation is not necessarily better than a short one.)

   (a) (True/False) Gauss-Newton method is always superior to data linearization since it does not distort the original equation. Your explanation should not be your opinion, but some remarks from textbook or notes.

   (b) (True/False) Gauss-Newton method, when a convergence was achieved, is always superior to data linearization when it comes to 2-norm of residual.

   (c) (True/False) Levenberg–Marquardt Method can show better conditioning behavior than Gauss-Newton method, but not necessarily a convergence to a better solution when there is a single minimizer of the 2-norm of residual and there is no local minimum.

   (d) (True/False) Data linearization and nonlinear least square method yield the same form of model, but with different parameters that determines the model.

5. (Formation) Answer the following.

   (a) List three approaches that we have learned to solving over-determined system of nonlinear equations. Also, give the main idea of each approach in one sentence.

   (b) Write out the Gauss-Newton method.

   (c) Write out the Levenberg–Marquardt Method.

   (d) Write the dot-product rule. If you prefer, use your favorite notation that is different from the textbook author. But, in that case, give a description of your notations, and possible why you like that notation.

   (e) Write the matrix-vector product rule. If you prefer, use your favorite notation that is different from the textbook author. But, in that case, give a description of your notations, and possible why you like that notation.

III. Only for presentation

The following problems are not graded, but available for presentation. The reason for exclusion is for you to spend enough time for the submitted problems in part I and part II so you really understand them down to the details.

6. (Exploration; Computation) Apply Levenberg-Marquardt to fit the model $y = c_1 e^{-c_2(t-c_3)^2}$ to the following data points, with an appropriate initial guess. State the initial guess, the regularization parameter λ used, and the RMSE. Plot the best least squares curve and the data points.

   $(t_i, y_i) = \{(1,1), (2,3), (4,7), (5,12), (6,13), (8,5), (9,2), (11,1)\}$

   End of homework