1 Introduction to the Project 2

The topic of this project is data-fitting. The main task is to find a fitting function $f(a, u)$ which approximates experimental data provided in data files

- `isobar_benzene_methanol.mat` which consist of $T - x - y$ equilibrium data for benzene-methanol mixture,
- `isobar_acetone_methanol.mat` which consist of $T - x - y$ equilibrium data for acetone-methanol mixture,
- `isothermal_acetone_methanol.mat` which consist of $p - x - y$ equilibrium data for acetone-methanol mixture,
- `isobar_ethanol_toluene.mat` which consist of $T - x - y$ equilibrium data for acetone-methanol mixture,
- `isothermal_ethanol_toluene.mat` which consist of $p - x - y$ equilibrium data for ethanol-toluene mixture.

Variable T stands for temperature in K, p for pressure in kPa, x and y are dimensionless molar fractions of substances in liquid and gaseous phases, respectively. The data fitting function has two arguments, the vector $a \in \mathbb{R}^n$ is the vector of optimized coefficients (see Eq.(2)) and u is the independent variable representing some physical quantity. Note, that for the purpose of fitting procedure, u is always known, it can be for example the molar fraction y of acetone.

Since there is no advance knowledge on the structure of the fitting function, it is necessary to test several function candidates and determine the best one. For such a purpose, consider following set of elementary functions

$$\mathcal{F} = \left\{ \mathrm{e}^{\alpha u},\ \mathrm{e}^{-\beta u},\ u,\ u^2,\ \frac{1}{u + 1.1},\ 1,\ \sqrt{u} \right\}, \tag{1}$$

from which you must choose some elementary functions to construct the candidate for $f(u)$, which may look like

$$f(a, u) = a_1 \cdot \mathrm{e}^{-\beta u} + a_2 \cdot u + a_3 \cdot 1. \tag{2}$$

Note, that you may choose multipliers $\alpha, \beta \in \mathbb{R}_+$, it may increase the goodness of the fit, but you may leave them set to $\alpha = \beta = 1$.

Once the candidate function is constructed, the optimization problem can be casted as follows

$$J^\star = \min_a \frac{1}{2m} \sum_{i=0}^{m} \left(v_i - f(a, u_i) \right)^2, \tag{3}$$

where m denotes the number of measurements and v_i is a specific measurement for a independent value of u_i. The value $J^\star$ serves as an indicator of the goodness of the fit. Naturally, for perfect fit $J^\star = 0$.

The optimization problem shall be solved with the Gradient-Descent Method (GDM) with backtracking and with the Newton method (NM) with backtracking. Convergence results shall be compared. Consider 4 candidate functions with combination of at least 3 elementary function from $\mathcal{F}$. Note, that you may extend the set $\mathcal{F}$ by your choice of elementary functions, such as higher polynomials or logarithmic functions.

2 Grading

Create a free repository in Mercurial or GIT, where you will track your progress. Link to the repository needs to be included in the MOODLE submission. The final submission must include the main file called `main.m`, functions `gdmb.m`, and `newtonb.m`. Modify GDM and NM scripts from exercises to accept an additional parameter tolerance, which can be set to 10^{-5} (for testing purposes use 10^{-2}).

The main file shall produce following set of figures

- *convergence figure*, demonstrating the convergence of GDM and NM for each of the candidate function (4×5p),
- *accuracy figure*, demonstrating the convergence of GDM and NM for each of the accuracy of the optimal solution (4×5p),
- *fitting figure*, consisting of data points and 4 curves representing the approximation via function candidate, in the title of the figure attach individual value of $J^\star$ (20p),
- *t-dependence figure*, demonstrating the number of iterations required to reach specified tolerance with varying step size t, such a figure is to be produced for only one optimization problem, which is solved with the GDM (20p).

Each student is required to optimize the source-code using anonymous function, for-loops, etc. Code optimization amounts to 10p, and 10p will be awarded for graphic presentation of results. Note, that submitted approximations which are obviously wrong will not be accepted. Furthermore, submitting something other then three m-files will result in failing this assignment.

3 Individual Assignments

Following table summarizes individual assignments:

student	mixture	approximation to be found
Diana Dzurková	Acetone-Methanol	$T = f(a, x)$
Kristína Fedorová	Acetone-Methanol	$T = f(a, y)$
Lenka Galčíková	Benzene-Methanol	$T = f(a, y)$
Samuel Hrstka	Acetone-Methanol	$p = f(a, x)$
Matej Kintler	Acetone-Methanol	$x = f(a, y)$
Marek Kišoň	Toluene-Ethanol	$T = f(a, x)$
Roman Kohút	Toluene-Ethanol	$T = f(a, y)$
Alexey Morozov	Toluene-Ethanol	$p = f(a, x)$
Michal Slávik	Toluene-Ethanol	$p = f(a, y)$
Rudolf Trautenberger	Benzene-Methanol	$T = f(a, x)$