BT 5240 — Computational Systems Biology

Jan–May 2020

Assignment 3

Due date: 4th April, 2020 @ 17:00 **Maximum marks: 60**

Academic Integrity: You are allowed to discuss the problems verbally with your friends, but copying or looking at codes (either from your friend or the Web) is not permitted. Transgressions are easy to find, and will be reported to the "Sub-committee for the Discipline and Welfare of Students" and will be dealt with very strictly. Mention any collaboration (discussions only!) in your solutions.
Late submission penalties: 1 second – 24 h: 20%; 24–48 h: 40%; $>$ 48h: 60%

Submission: Since this is a computational assignment, I would also like to look at your codes. Submit your assignment as **one zip file** by uploading it at http://tinyurl.com/bt5240-submit. Your zip file should be named something like `BTyyBxxx.zip`, based on your roll number. This zip file must contain a single neatly typeset PDF of your solutions (named `BTyyBxxx.pdf`) as well as the codes used for each of problems in a separate folder `codes`.

Problem Statement A model of the MAPK cascade is defined by the kinetic equations detailed below:

$$\frac{d[MKKK]}{dt} = r_2 - r_1$$
$$\frac{d[MKKK-P]}{dt} = r_1 - r_2$$
$$\frac{d[MKK]}{dt} = r_6 - r_3$$
$$\frac{d[MKK-P]}{dt} = r_3 + r_5 - r_4 - r_6$$
$$\frac{d[MKK-PP]}{dt} = r_4 - r_5$$
$$\frac{d[MAPK]}{dt} = r_{10} - r_7$$
$$\frac{d[MAPK-P]}{dt} = r_7 + r_9 - r_8 - r_{10}$$
$$\frac{d[MAPK-PP]}{dt} = r_8 - r_9$$

Reactions:

$$r_1 = V_1 * \frac{[MKKK]}{((1 + (\frac{[MAPK-PP]}{KI})^n) * (K_1 + [MKKK]))}$$

$$r_2 = p_3 * \frac{[MKKK-P]}{(K_2 + [MKKK-P])}$$

$$r_3 = p_1 * [MKKK-P] * \frac{[MKK]}{(K_3 + [MKK])}; \quad r_4 = p_1 * [MKKK-P] * \frac{[MKK-P]}{(K_4 + [MKK-P])}$$

$$r_5 = p_2 * \frac{[MKK-PP]}{(K_5 + [MKK-PP])}; \quad r_6 = p_2 * \frac{[MKK-P]}{(K_6 + [MKK-P])}$$

$$r_7 = p_1 * [MKK-PP] * \frac{[MAPK]}{(K_7 + [MAPK])}; \quad r_8 = p_1 * [MKK-PP] * \frac{[MAPK-P]}{(K_8 + [MAPK-P])}$$

$$r_9 = V_9 * \frac{[MAPK-PP]}{(K_9 + [MAPK-PP])}; \quad r_{10} = V_{10} * \frac{[MAPK-P]}{(K_{10} + [MAPK-P])}$$

Moiety conservation relations:
$[MKKK]_{total} = [MKKK] + [MKKK-P] = 100$
$[MKK]_{total} = [MKK] + [MKK-P] + [MKK-PP] = 300$
$[MAPK]_{total} = [MAPK] + [MAPK-P] + [MAPK-PP] = 300$

Parameters:
Set 1: $K_I = 9, n = 1, V_1 = 2.5, V_9 = V_{10} = 0.5, K_1 = 10, K_2 = 8, K_3 = K_4 = K_5 = K_6 = K_7 = K_8 = K_9 = K_{10} = 15$.
Set 2: $K_I = 18, n = 2, V_1 = 2.5, V_9 = V_{10} = 1.25, K_1 = 50, K_2 = 40, K_3 = K_4 = K_5 = K_6 = K_7 = K_8 = K_9 = K_{10} = 100$.

Assume that we have experimental data of concentrations of the various species at different time-points. The experimental data is noisy and there are some time-points at which data could not be recorded for certain species. We have **two data sets** D_1 and D_2 which have concentration data for the different parameter sets mentioned above. **Estimate the parameters** p_1**,** p_2 **and** p_3, and comment on the squared sum of errors for the different parameter sets.
For the ODE simulations use the initial conditions:
$[MKKK] = 90, [MKKK-P] = 10, [MKK] = 280, [MKK-P] = 10, [MKK-PP] = 10, [MAPK] = 280, [MAPK-P] = 10, [MAPK-PP] = 10$

Use any algorithm discussed in class (*e. g.* standard minimization algorithm like `fminsearch` or `fmincon` in MATLAB/Simulated annealing/Evolutionary algorithms) — no need to write your own code (use publicly available libraries).

HINT: Constrain your search space using a lower bound of 0 and an upper bound of 10.

Repeat the above for three combinations of the data sets: D_1, D_2 and $D_1 \cup D_2$, and comment.