Evolutionary Estimation of Parameters in Computational Models of Thymocyte Dynamics

Lavinia Moatar-Moleriu, Viorel Negru, and Daniela Zaharie

Department of Computer Science, West University of Timişoara, Romania
{lmoatar,vnegru,dzaharie}@info.uvt.ro

Abstract. This paper presents an evolutionary-based parameter estimation procedure able to deal with the particularities of the constraints arising in mathematical models of biological systems. A measure of the constraint satisfaction degree and several feasibility-based ranking rules are proposed and comparatively analyzed for the problem of estimating the parameters involved in a model describing the dynamics of thymocytes. The numerical results illustrate the effectiveness of the procedure in inferring models which fit well the experimental data and also satisfy the biological constraints.

Keywords: population dynamics, parameter estimation, constrained optimization, differential evolution

1 Introduction

Inferring models from experimental data is an important task in computational biology and usually leads to difficult constrained optimization problems. As the estimation of the model quality can involve simulation of complex systems and the apriori knowledge on the parameters could be limited, local optimization methods involving gradient computation are inapplicable. Therefore, metaheuristics proved to be viable methods for parameter estimation of such models [7]. In this paper we focus on the problem of identifying computational models able to simulate the dynamics of thymocyte populations taking place in the thymus, as part of the complex process through which the organisms defend against infections [2]. Aiming to model transient perturbations of the normal dynamics of thymocytes we arrived to the problem of estimating several dozens of parameters such that some biologically motivated constraints are satisfied. As these constraints are related to properties of some time-dependent functions they require specific handling techniques. The paper is organized as follows. Section 2 presents the mathematical model and the components of the constrained optimization problem. Section 3 shortly reviews evolutionary constraint optimization while the proposed parameter estimation procedure, including the specific constraint handling variants, is presented in Section 4. Results of a comparative analysis and the numerical validation of the estimation procedure are presented in Section 5. Finally, Section 6 concludes the paper.

2 The mathematical model and the parameter estimation problem

One of the simplest models describing the dynamics of thymocyte populations, proposed in [2], consists of four coupled differential equations (Eqs. (1)). Each equation describes the evolution of the number of cells in the corresponding population, controlled by proliferation, death and transfer rates (denoted by r, d and s, respectively). Besides these rates, the model contains other three parameters: b (inflow of progenitor cells), K and K_n (carrying capacities).

$$\begin{aligned}
\dot{N}(t) &= r_n N(t)(1 - N(t)/K_n) - d_n N(t) - s_n N(t) + b(1 - N(t)/K_n) \\
\dot{P}(t) &= r_p P(t)(1 - Z(t)/K) - d_p P(t) - (s_4 + s_8)P(t) + s_n N(t) \\
\dot{M}_4(t) &= r_4 M_4(t)(1 - Z(t)/K) - d_4 M_4(t) - s_{o4} M_4(t) + s_4 P(t) \\
\dot{M}_8(t) &= r_8 M_8(t)(1 - Z(t)/K) - d_8 M_8(t) - s_{o8} M_8(t) + s_8 P(t) \\
Z(t) &= N(t) + P(t) + M_4(t) + M_8(t)
\end{aligned} \tag{1}$$

These equations proved to be appropriate in modelling the thymocyte dynamics in a normal thymus [2]. However various pathological situations or the administration of some substances can perturb the normal dynamics by inducing a significant involution followed by a regeneration of the thymocyte populations. Such a dynamics can be simulated by replacing the constant rate parameters in Eqs.(1) with variable rates obtained by adding to the intial rates a time-depending function which models the transient perturbation. A family of functions appropriate to model a perturbation starting from a zero value at an initial time moment and approaching again zero after a time interval is described in Eq.(2), where $C = \{c_1, c_2, c_3, c_4, c_5\}$ denotes a set of positive parameters.

$$\xi(C;t) = \frac{c_1}{t^{c_3} + c_2} - \frac{c_1 c_4 / c_2}{t^{c_5} + c_4} \tag{2}$$

By replacing each constant rate r with $r + \xi(C;t)$, five new parameters are introduced for each of the thirteen rates, leading to a set of $k = 71$ parameters in the model. Estimating the parameters values means finding $x^* \in \mathbf{R}^k$ which minimizes the mean squared error described in Eq.(3) and satisfies constraints related to the positivity of all perturbed rates and to the vanishing of the perturbation.

$$MSE(x) = \frac{1}{4n} \sum_{\pi \in \{N,P,M_4,M_8\}} \left(\frac{1}{\max_{j=\overline{1,n}} \{\bar{\pi}_j^2\}} \sum_{j=1}^{n} (\pi(x;t_j) - \bar{\pi}_j)^2 \right) \tag{3}$$

In Eq.(3) n denotes the number of experimental values available for each thymocyte population, $\bar{\pi}$ denotes experimental values corresponding to each of the four populations and $\pi(x;t)$ denote numerically estimated solutions corresponding to the given set of parameters and to the time moments of the experimental measurements. These estimated solutions are obtained by numerically solving Eqs.(1) for initial values compatible with the experimental data. The division of the error terms corresponding to each population by the maximal measured value ensures the balance between the errors corresponding to different thymocyte populations and avoid the bias in the estimation process toward parameters of the dominant population. For each perturbed rate the two constraints to be satisfied are described in Eq. (4) where t_a denotes the time moment when the perturbation starts and t_f the time moment when it should be small enough (e.g. smaller than a given value $\epsilon_f > 0$).

$$r + \xi(C;t) \geq 0 \text{ for all } t \in [t_a, t_f]; \quad |\xi(C;t_f)| < \epsilon_f \tag{4}$$

3 Evolutionary constrained optimization

A constrained optimization problem usually requires to find $x^* \in \mathbf{R}^k$ which minimizes an objective function $f : \mathbf{R}^k \to \mathbf{R}$ and satisfies $g_j(x^*) \geq 0$ for each $j \in \{1, \ldots, q\}$. Evolutionary constrained optimization relies on using a constraint handling technique when applying the basic evolutionary operators (variation and/or selection). Most approaches interferes with the selection process by changing either the fitness value computation (penalty method) or the comparison rule between two candidate solutions (feasibility rules, stochastic ranking) [3]. Despite the differences between them, all these techniques uses the so-called constraint violation amount which for a constraint $g_j(x) \geq 0$ is defined by $\phi_j(x) = \min(0, g_j(x))^2$. The overall violation of the constraints is defined as the sum $\phi(x) = \phi_1(x) + \ldots + \phi_q(x)$. While in the penalty function technique the objective function and the constraint violation function are combined, in the feasibility based rules they are separately used. The classical Deb's rule [1] specifies that a candidate solution x is better than x' if one of following conditions is satisfied: *(i)* x is feasible and x' is not feasible; *(ii)* both x and x' are feasible and $f(x) < f(x')$; *(iii)* both x and x' are infeasible and $\phi(x) < \phi(x')$.

Using the objective function as comparison criterion only in the case of feasible solutions can lead to premature convergence [3]. Two variants which enlarges the set of cases when the objective function is used as optimization criterion are ϵ-feasibility [6] and stochastic ranking [5]. The ϵ-feasibility rule is based on a relaxation of the feasibility concept, i.e. x is considered better than x' if one of following conditions is satisfied: *(i)* $\phi(x) \leq \epsilon$ and $\phi(x') \leq \epsilon$ (x and x' are almost feasible) and $f(x) < f(x')$; *(ii)* $\phi(x) = \phi(x')$ (same constraints violation) and $f(x) < f(x')$; *(iii)* $\phi(x) < \phi(x')$. On the other hand, the stochastic ranking enhances the role of the objective function by involving it in the decision rule not only when the solutions are feasible but also when a random event occurs. However all these feasibility based rules use, when deciding if a solution is better than another, the objective function and the constraints violation function not in an aggregated but in a separate way.

4 The proposed parameter estimation procedure

The parameters of the model described in Section 2 which should be estimated can be grouped in several sets: $(C_0; C_1; \ldots; C_q)$. C_0 denotes the parameters involved in Eqs. (1) (i.e. $C_0 = (r_n,\, r_p,\, r_4,\, r_8,\, d_n,\, d_p,\, d_4,\, d_8,\, s_n,\, s_4,\, s_8,\, s_{o4},\, s_{o8},\, b,\, K_n,\, K))$. The sets C_j correspond to the parameters involved in the functions used to perturb the rates specified in C_0 and should satisfy the constraints described in Eq.(4). The first constraint is particularly difficult to check as, in the general case, it requires the analysis of the values of $r + \xi(C;t)$ over a continuous time interval. As a consequence it is neither easy to decide if a solution is feasible nor to compute the constraint violation amount. However for smooth continuous functions $\xi(C;t)$ one can estimate the degree of constraint satisfaction by sampling the time interval and computing the proportion of cases when the constraint is satisfied. More specifically, by considering an uniform discretization $T_h = \{t_a, t_a + h, t_a + 2h, \ldots, t_f\}$ of $[t_a, t_f]$ we can compute an estimation of the constraint satisfaction degree as given in Eq.(5).

$$S_p^j(C_j) = \frac{\text{card}\{t \in T_h | r_j + \xi(C_j;t) > 0\} - \delta}{\text{card}(T_h)} \tag{5}$$

The constant $\delta > 0$ in Eq.(5) has a small value and is used only to discriminate the cases when the constraint satisfaction can be mathematically proved. For instance in the particular case of constraints given in Eq. (4) a sufficient condition ensuring the positivity is $r \geq \max\{c_1/(c_2 + c_2^2), c_1 c_4/(c_2 + c_2 c_4)\}$. Therefore this condition is first checked and if it is satisfied then the positivity constraint is considered satisfied and S_p is set to 1. Otherwise S_p is computed using Eq. (5).

For the second type of constraints the satisfaction degree can be computed following a standard approach which leads to a value S_v as defined in Eq.(6).

$$S_v^j(C_j) = \begin{cases} 1 & \text{if } |\xi(C_j;t_f)| \leq \epsilon_f \\ 1 - \min\{1, |\xi(C_j;t_f)|\} & \text{otherwise} \end{cases} \tag{6}$$

The overall degree of constraints satisfaction, $S \in [0,1]$ is defined as $S(C) = \prod_{j=1}^{q} S_p^j(C_j) S_v^j(C_j)$. The values of S can be interpreted as follows: *(i)* if $S(C) = 1$ then C is surely feasible; *(ii)* if $S(C) \geq 1 - \delta/\text{card}(T_h)$ then C is probably feasible (there is neither evidence that the first constraint is violated nor guarantees that it is satisfied); *(iii)* if $S(C) = 0$ then at least one of the constraints is severely violated (at least one perturbation is too large or for at least one perturbed rate all sampled values are negative); *(iv)* in all other cases, the value of S offers information about the degree of constraint satisfaction.

Examples of several cases of perturbed rates satisfying or violating the constraints and the corresponding S values are illustrated in Fig. 1. Having a value in $[0,1]$, S can be used to penalize the value of the objective function or as acceptance probability of infeasible configurations.

Constraints handling. There are several ways to use the satisfaction degree S and the MSE value in order to decide which of two candidate solutions is better. Starting from the existing feasibility and ranking rules [3] and using the properties of S we identified several variants which we comparatively analyzed with respect to their effectiveness in solving the addressed parameter estimation problem.

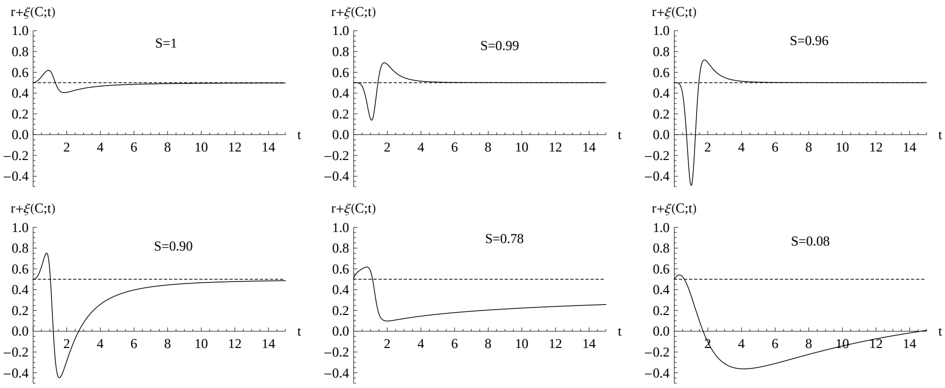

Fig. 1. Illustration of the relationship between the properties of the perturbed rates and the values of the constraint satisfaction degree, S. Continuous line: perturbed rate, dashed line: initial value of the rate.

Ranking rule A. Using the assumption that $S(x) \geq \theta$ suggests that x is feasible, while $S(x) < \theta$ means that it is infeasible (for a given threshold θ), the Deb's feasibility rule can rewritten as follows. A candidate solution x is better than x' if one of the following conditions is satisfied: *(i)* $S(x) \geq \theta$ and $S(x') < \theta$; *(ii)* $S(x) \geq \theta$ and $S(x') \geq \theta$ and $MSE(x) < MSE(x')$; *(iii)* $S(x) < \theta$ and $S(x') < \theta$ and $S(x) \geq S(x')$.

Ranking rule B. One of the particularities of the previous rule is that the objective function and the constraint satisfaction degree are used in a decoupled way. A variant which aggregates MSE and S states that x is better than x' if one of the following conditions is satisfied (checked in this specific order): *(i)* $S(x) \geq \theta$ and $S(x') < \theta$; *(ii)* $S(x)S(x') = 0$ and $MSE(x) \leq MSE(x')$; *(iii)* $S(x) \neq 0$ and $S(x') \neq 0$ and $MSE(x)/S(x) \leq MSE(x')/S(x')$.

Ranking rule C. The first two rules analyze first the cases when the constraints are satisfied or close to be satisfied. By ruling out first the cases when the constraints are severely violated one obtains a slightly different variant when x is better than x' if: *(i)* $S(x) > 0$ and $S(x') = 0$; *(ii)* $S(x) = 0$ and $S(x') = 0$ and $MSE(x) \leq MSE(x')$; *(iii)* $S(x) \neq 0$ and $S(x') \neq 0$ and $MSE(x)/S(x) \leq MSE(x')/S(x')$.

Ranking rule D. Instead of inferring the feasibility in a deterministic way from the value of S one can do it in a probabilistic manner. In this case $S(x)$ is interpreted as a probability that x, if selected, can lead to a feasible configuration. Thus, by denoting with U_1 and U_2 two independent random values uniformly selected from $[0, 1]$ one can say that x is better than x' if one of the following conditions is satisfied: *(i)* $U_1 \leq S(x)$ and $U_2 > S(x')$; *(ii)* $U_1 > S(x)$ and $U_2 > S(x')$ and $MSE(x) \leq MSE(x')$; *(iii)* $U_1 \leq S(x)$ and $U_2 \leq S(x')$ and $MSE(x)/S(x) \leq MSE(x')/S(x')$.

Ranking rule E. Starting from the idea of stochastic ranking [5], which allows in a probabilistic manner to use the objective function as comparison criterion, even for infeasible solutions, we arrived at the following rule which states that x is better than x' by sequentially checking the following conditions: *(i)* $S(x) \geq \theta$, $S(x') \geq \theta$ and $MSE(x) < MSE(x')$; *(ii)* $U < P_f$, $S(x)S(x') \neq 0$ and $MSE(x)/S(x) < MSE(x')/S(x')$; *(iii)* $S(x) \geq S(x')$.

Search method. As stated in [3] one of the most competitive evolutionary metaheuristic in solving constrained optimization problems seems to be Differential Evolution (DE). On the other hand the effectiveness of DE for parameter estimation in biological systems was reported in several comparative studies [7]. This motivated us to use JADE, an adaptive DE variant introduced in [8]. The JADE overall structure is described in Algorithm 1 and its main features are: (i) the elements used in the recombination rule described in Eq. (7) are chosen such that a new candidate is created in a neighborhood of a good element but away from a worse one (x_{rbest} is selected from the $p\%$ elites of the current population and x_{r2} is one of the inferior elements which were discarded in a previous selection step and was stored in an archive); (ii) the scale factor (F) and the crossover probability (CR) are generated for each element of the population using a probability distribution (Gaussian and Cauchy, respectively) whose mean is recomputed at each generation using information from successful elements.

$$z_i^l = \begin{cases} x_i^l + F_k \cdot (x_{rbest}^l - x_i^l) + F_i \cdot (x_{r1}^l - x_{r2}^l) & \text{if } rand() \leq CR_i \\ x_i^l & \text{otherwise} \end{cases}, i = \overline{1,m}, j = \overline{1,k} \tag{7}$$

The constraint handling techniques interfere with two of the JADE components: (i) the ranking process used to select the top $p\%$ elements; (ii) the selection of the survivor between the parent and the trial element.

Algorithm 1 JADE overall structure

```
Initialization step (population, control parameters, archive)
while ⟨the stopping condition is false⟩ do
    Rank the population and identify the top p% elements
    for i = 1,m do
        Construct z_i using Eq.(7); Choose the best between z_i and x_i
    end for
    Update the control parameters and the archive
end while
```

5 Numerical results

The used experimental dataset consists of 232 estimates of the number of cells in each of the four thymocyte populations collected from young and adult mice thymus either before or after a treatment administration. Each of the five ranking rules proposed in the previous section was combined with JADE leading to a specific procedure to estimate all $k = 71$ parameters of the model which satisfies $q = 26$ constraints. In each case, results from 30 independent runs were collected. The results reported in Table 1 have been obtained by using populations of $m = 20$ elements, 5000 generations, a percent $p = 10\%$ in defining the set of top elements and 0.5 as initial mean of the distributions used to provide values for F and CR. A preliminary numerical study using several population sizes (e.g. 20, 50 and 100) suggested that $m = 20$ leads to the best quality/cost trade-off. For the variant inspired by stochastic ranking, the value of P_f was set to 0.45, as suggested in [5]. The numerical solutions of the system (1) required for MSE estimation were obtained using the ODE solver from Mathematica 7.0.

Table 1 presents statistical values of MSE, the constrained satisfaction degree S (with $\epsilon_f = \delta = 0.001$) and the feasibility probability (FP) defined as the ratio between the number of runs when the solution can be considered feasible (i.e. $S \geq \theta$) and the total number of runs [4]. As feasibility threshold, any value θ larger than 0.99 proved to lead to solutions satisfying the positivity constraint. The value of the threshold corresponding to the cases when all sampled values of the perturbed rates are positive is $1 - \delta/\text{card}(T_h) = 0.999927$ (for $h = 0.1$ and $\delta = 0.001$). The results in Table 1 show that the analyzed ranking rules are characterized by different quality of fit vs. constraint satisfaction trade-offs. The best trade-off is obtained by the variant inspired from stochastic ranking which with respect to MSE is superior to other ranking rules. The p-values obtained when a Mann-Whitney test was applied to compare rule E and the other rules are reported in the last column of Table 1. With respect to the constraint satisfaction the most effective ones are rules B and D. Best behavior was observed for the ranking rules using an aggregation of MSE and S. The ability of the proposed procedure to lead to a model which fits well to the data and satisfies the constraints on rates is illustrated (for one of the four populations) in Fig. 2.

Table 1. Quality of fit (MSE), constraints satisfaction (S), feasibility probability (FP, $\theta_1 = 0.999927$ and $\theta_2 = 0.99$) and p-values (Mann-Whitney test for comparing rule E and the other rules). Algorithm: JADE combined with the proposed ranking rules.

Ranking rule	MSE	S	$FP(\theta_1)$	$FP(\theta_2)$	p-value
Rule A ($\theta = 1$)	0.0338 ± 0.0012	1 ± 0	1	1	$5 \cdot 10^{-12}$
Rule A ($\theta = 0.99$)	0.0270 ± 0.0010	0.9966 ± 0.0033	0.5	1	$7 \cdot 10^{-8}$
Rule B ($\theta = 0.99$)	0.0268 ± 0.0014	$0.9999 \pm 5 \cdot 10^{-6}$	1	1	$3 \cdot 10^{-6}$
Rule C	0.0261 ± 0.0009	0.9878 ± 0.0119	0.45	0.45	$4 \cdot 10^{-5}$
Rule D	0.0290 ± 0.0017	$0.9999 \pm 3 \cdot 10^{-6}$	1	1	$6 \cdot 10^{-10}$
Rule E ($P_f = 0.45$)	0.0250 ± 0.0005	0.9935 ± 0.0011	0.03	1	-
Unconstrained	0.0208 ± 0.0022	0.0468 ± 0.0776	0	0	-

6 Conclusions

By combining an evolutionary algorithm with an appropriate constraint handling technique we succeeded in inferring a model of the perturbed thymus dynamics which is in accordance with the experimental data. The constraint satisfaction

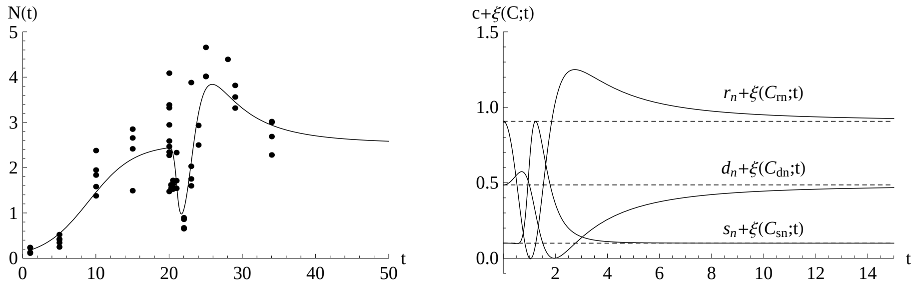

Fig. 2. Left: Experimental data and simulated dynamics of N ($MSE = 0.023$, $S = 0.993$, $t_a = 20$, $t_f = 35$ days). Right: Initial rates (dashed lines) and perturbed rates (continuous lines) for N.

degree and the proposed ranking rules, characterized by aggregating the quality of fit measure and the constraint satisfaction degree, can be used for other optimization problems involving constraints which can be only partially checked.

Acknowledgments. This work was supported by grant no. PN-II-ID-PCE-2011-3-0571 and by the strategic grant POSDRU/CPP107/DMI1.5/S/78421, Project ID 78421 (2010), co-financed by the European Social Fund Investing in People, within the Sectoral Operational Programme Human Resources Development 2007-2013. The authors thank dr. Felix Mic (University of Medicine and Pharmacy, Timisoara) for providing the experimental data.

References

1. Deb, K.: An efficient constraint handling method for genetic algorithms. Computer Methods in Applied Mechanics and Engineering. 186, 311-338 (2000)
2. Mehr, R., Globerson, A., Perelson, A.S.: Modeling positive and negative selection and differentiation processes in the thymus. J. Theor. Biol. 175, 103–126 (1995)
3. Mezura-Montes, E., Coello Coello, C.A.: Constraint-handling in nature-inspired numerical optimization: Past, present and future. Swarm and Evolutionary Computation, 1, 173–194 (2011)
4. Mezura-Montes, E., Miranda-Varela, M.E., Gomez-Ramon, R.C.: Differential evolution in constrained numerical optimization: An empirical study. Information Sciences, 180, 4223–4262 (2010)
5. Runarsson,T.P., Yao, X.: Stochastic ranking for constrained evolutionary optimization. IEEE Transactions on Evolutionary Computation, 4 284-294 (2000)
6. Takahama, T., Sakai, S.: Constrained optimization by ϵ-constrained particle swarm optimizer with ϵ-level control. In Proc. of WSTST'05, 1019–1029 (2005)
7. Tashkova, K., Korošek,P., Šilc,J., Todorovski,L., Džeroski,S.: Parameter estimation with bio-inspired meta-heuristic optimization: modeling the dynamics of endocytosis. BMC Systems Biology, 5:159, doi:10.1186/1752-0509-5-159 (2011)
8. Zhang, J., Sanderson, A.C.: JADE: adaptive differential evolution with optional external archive. IEEE Trans. on Evolutionary Computation, 13(5), 945–958 (2009)