Neighborhood Search Considering Continuous Space Search

A.T. GÓMEZ, B.E.J. BODMANN, Centro de Ciências Exatas e Tecnológicas, Universidade do Vale do Rio dos Sinos–UNISINOS, Av. Unisinos 950, 93022-000 São Leopoldo, RS, Brasil.

Abstract. The objective of this work is to present and discuss a dynamical optimization procedure based on the algorithm of self-consistent parametric inference, where the movement of one feasible solution to another is implemented in continuous mode. The approach is based on a generic parametrization of search space and uses the definition of a model independent metric, which guides progress of the procedure. One of the principal contributions of the proposed method is, if a solution is found it is optimal for the parametrization considered. This feature is guaranteed and proven by self consistency.

1. Introduction

Usually, solving procedures in non-trivial nonlinear optimization problems, such as resource allocation, scheduling and storage among others, are classified as NP-hard problems. In order to render those challenges more tractable, from the last decade onwards, approaches based on linear and integer programming have been successively replaced by contemporary heuristics techniques as for instance simulated annealing, genetic algorithms and tabu search [1, 2]. Typically the structure of those problems show complex and highly non-linear interrelations between its descriptors so that a traditional analytic approach is limited by always necessary approximations involved and by restrictions to non-stochastic quantities.

The scope of the present work is to discuss a heuristic technique, that permits an analytic description which may have stochastic character and additionally is reasonably free of inherent model assumptions, i.e. has got a generic character, except for the necessary parametrization. As will be shown in section 4., the parametrization is strongly related to the model variables and in this sense belongs to the consistent system description rather than to the employed optimization procedure. The obtained functional representation permits a global analysis in continuous space search as lined out in detail in [3]. This type of approach gains importance once a rigorous proof for the optimal or close to optimal solution is of need, which is especially of interest in cases where heuristic knowledge or experience is vague, and decisions may be based on hypothesis testing only. Note, that the present method breaks with a widely accepted view, that analyticity and stochastic features are hardly to be implemented in one unique procedure [4].

2.   The master equation

Without restrictions, a variational or differential problem may always be cast into integral form especially once constraints or boundary conditions are present [5], which defines the starting point for the proposed self consistency treatment. Let $\{T_i\}$ be a set of decision variables and t an independent variable. The variable interrelation is then defined by a set of differential equations

$$\frac{\partial T_i}{\partial t} = \sum_\alpha \prod_i c_{i\alpha} T_i^{\alpha_i} \ , \qquad (2.1)$$

where $c_{i\alpha}$ are coefficients to be determined by optimization and $\alpha \in I\!R^N$ a multi-index. To justify this conjecture, it is a well established experience in dynamical programming that in general problem solvers are more likely to be defined via self consistent equations [6, 7], that is, the solution itself affects its local behavior. Equation (2.1) defines a network like structure with weight factors $c_{i\alpha}$ at each vertex. Upon integration and summing all equations involved yields the master equation

$$T_i = T_i^{(ini)} + \int_{\tau_0}^{\tau} f(\{T_i\}|t) \ dt \ . \qquad (2.2)$$

The above equation is manifestly exact. However, in probably almost all cases of interest there is lack of an analytical solution for T_i, which is the main issue of the present work. The integral plays the role of an "objective function" with the decision variables setup in a self consistent fashion. We then want to obtain estimates of the unknown parameters $c_{i\alpha}$, hence our question about the correlation between the trial function and the true function is put in terms of the model – well established as parametric inference, with $\{c_{i\alpha}\}$ to be determined by measuring self consistency of the underlying parameter approach.

3.   The self consistency approach

Our reasoning for establishing the self consistency approach is as follows: On substitution of T_i by the parametrization $T_i(\{c_{i\alpha}\}|t)$, the integral kernel (i.e. formalized heuristics) $f(\{T_i\}|t)$) may be interpreted as a "functional sample" of infinite size $(\tau - \tau_0)/dt$, with a distribution of $\{c_{i\alpha}\}$s, whose function to be determined to a reasonable approximation is $T_i(\{c_{i\alpha}\}|\tau)$, as defined above. To be more specific, the kernel f composes the solution (left hand side of (2.2)) as a continuous sequence of $f(T_i)$s at each coordinate point $t \in [\tau_0, \tau]$, with the set of $\{c_{i\alpha}\}$ to be adjusted at each t to establish point-wise equality of the equation. In order to apply parametric inference techniques, a deviation measure for the "functional sample" must be defined, i.e. the self consistency measure $\mathcal{S}$. Since the real solution is substituted by a known trial function a self consistency measure may be constructed, in close analogy to the Likelihood function [8]. Because of the functional sample being of infinite size, this function is defined as the limit of a regularized product, which may

be recognized as the consistency for observing the continuous sequence $T_i(\{c_{i\alpha}\}|t)$ with $t \in [\tau_0, \tau]$.

$$\mathcal{S}(\{c_{i\alpha}\}|\tau) = \lim_{n\to\infty} \left\{ \prod_{\lambda=1}^{n} \mathcal{D}(T_i(\{c_{i\alpha}\}|t_\lambda) \right\}^{\Delta t} = \exp\left\{ \int_{\tau_0}^{\tau} \ln \mathcal{D} \, dt \right\} ,$$

where $t_\lambda = (1-\frac{\lambda}{n})\tau_0 + \frac{\lambda}{n}\tau$, and $\Delta t = (\tau-\tau_0)/n$, and $\mathcal{D}$ a self consistency generator to measure the inverse of the deviation of the true solution T_i from the parametrization $T_i(\{c_{i\alpha}\}|t)$ with the following properties:

1. In the limit $T_i(\{c_{i\alpha}\}|t) \to T_i \ \forall t \in [\tau_0, \tau], \ \mathcal{D} \to \infty$.

2. If $T_i(\{c_{i\alpha}\}|t) \neq T_i \ \wedge \ t \in [\tau_0, \tau], \ \mathcal{D} > 0$.

In order to obtain an acceptable solution to the problem the function $\mathcal{S}$ has to show a maximum self consistency for the trial function $T_i(\{c_{i\alpha}\}|\tau)$, i.e. the set of parameters must be such that $\mathcal{S}$ is a maximum. Assume now that in the range of interest, $\mathcal{S}$ be a well behaved function, i.e. regular, then the above condition is true if the set of equations

$$\left\{ \frac{\partial \ln \mathcal{S}}{\partial c_{i\alpha}} = 0 \right\} \tag{3.1}$$

has a solution. Equation (3.1) yields then the maximum self consistency estimators denominated the maximum self consistency parameter set $\{\tilde{c}_{i\alpha}\}$.

It remains to determine the quality of the *ansatz* for the self consistency problem. To this end one has to determine the variance of the consistency parameter set distribution, which may be done using the proper self consistency function. Recalling, that $\mathcal{S}$ represents the consistency for reproducing the continuous sequence $T_i(\{c_{i\alpha}\}|\tau)$, one may use $\mathcal{S}$ to calculate the distribution for $\{c_{i\alpha}\}$ and its variance

$$\sigma^2(\{c_{i\alpha}\}) = \int (\tilde{c}_{i\alpha} - c_{i\alpha})^2 \mathcal{S}(T_i(\{c_{i\alpha}\}|\tau)) \, [dt], \tag{3.2}$$

with $\tilde{c}_{i\alpha}$ the best estimator and $[dt]$ a regularized integral measure [5], which is understood as a result of the infinite "sample size" $(\tau - \tau_0)/dt$. Since this result is hardly obtainable even using numerical methods one may instead calculate the inverse of second-order partial derivative matrix

$$\mathcal{J}_{ij} = -\frac{\partial^2 \ln \mathcal{S}}{\partial c_{i\alpha} \partial c_{j\alpha}} , \tag{3.3}$$

where the diagonal elements are related to σ^2 via

$$\sigma_i^2(\{\tilde{c}_{i\alpha}\}) \approx \left(\mathcal{J}^{-1} \right)_{ii} , \tag{3.4}$$

a relation proved by Cramér and Rao [9]. So far only the formal procedure of how to solve for the unknown parameters was discussed, the next section is dedicated to the task of transforming the master equation into the generator $\mathcal{D}$ according to definition (4.2).

4.   The self consistency generator

It is now in order to generate an explicit self consistency measure from eq. (2.2). Recalling that the self consistency function must be well behaved, $\mathcal{D}$ must be positive definite and for almost all cases will be finite. An obvious construction, nevertheless not the only alternative, is making use of the $2n$-th moment $\mathcal{M}^{2n}$, where $n \geq 1$ is to be determined by convenience. For instance, in some cases it may be useful to take the limit $n \rightarrow \infty$ to obtain a critical boundary constraint on the parameter domain

$$\mathcal{M}^{2n} = \int_{\tau_0}^{\tau} \left(\frac{\partial T_i(\{c_{i\alpha}\}|t)}{\partial t} - f(T_i(\{c_{i\alpha}\}|t)) \right)^{2n} dt \ . \tag{4.1}$$

Note that, for $n = 1$ and $c_{i\alpha} \rightarrow \tilde{c}_{i\alpha}$, $\mathcal{M}^{2n}$ is equal to the total variance σ^2. We define the distribution as a Gaussian expression, using the second moment and the variance,

$$\mathcal{D}(\{c_{i\alpha}\}|\tau) = \frac{1}{\sqrt[n]{2\pi}\sigma} \exp \left\{ -\frac{\mathcal{M}^{2n}}{2n\sigma^{2n}} \right\} \ , \tag{4.2}$$

which complies with the characteristics of the desired self consistency measure. So far we elaborated the principal part of the procedure containing all necessary information to apply the method to an optimization problem. However, for an explicit application to already understood problems of operational research, we refer to a subsequent work (in preparation). In the following section we sketch out the algorithm structure which embeds the process of parameter estimation shown in previous sections and the logistics that will lead to a progressive improvement of the solution.

5.   The algorithm

Having elaborated the formal prerequisites it is now in order to design the algorithm, which is to assign specific values to the variables and solve the (in general) non-linear eigenvalue problem for the eigenvector (objective function) T_i with corresponding eigenvalue 1 in the parameter hyperspace. The solving procedure then reduces the initial manifold of dimension $(n+1)$ to a trajectory, where n denotes the dimension of the parameter space and "1" is due to the trajectory variable, respectively. The use of a manifold rather than approaches adopted in algorithms of the linear and non-linear type, discrete dynamical optimization, graph-theory or game theory based methods, combinatorial algorithms and hybrid methods is justified by the objective of developing a continuous space approximation method instead of a discrete one as the ones cited above (see for example [10, 11, 12, 13, 14, 15, 16, 17]).

Let $g(\{c_{i\alpha}\}|t)$ be a constraint with existing integrable differential form

$$dg(\{c_{i\alpha}\}|t) = 0, \ \forall t \in [\tau_0, \tau] \ ,$$

then $g(\{c_{i\alpha}\}|t)$ is called a holonomic constraint. Any system which is subject to one or more holonomic constraints is called a holonomic system, otherwise it is called a

non-holonomic system. Holonomic constraints by virtue may be included into the integral equation (2.2) making use of the Lagrange multiplier method [19]. Note, that non-holonomic constraints of the problem are typically present as inequalities.

It is the task of the solving procedure to decode the local information with the objective to derive global properties of a system under consideration. Lets suppose that a problem is setup by an integral equation of the type (2.2) with given kernel f and an initial boundary condition. Then there may or may not exist a solution to the specified problem, which will turn out as a result of the self consistency test eq. (3.1). The respective flow chart diagram of the procedure is shown in figure 1. Given an initial condition, if there exists a formal solution to the problem then, according to the theorem of Picard and Lindelöf [18], there exists a unique solution to the problem. Since in a bounded domain every compact solution which is not constant has a finite number of extremal points, it is guaranteed that the global optimum may be found. Whenever there appears the case of a degeneracy, i.e. the existence of more than one solution with equivalent optima, then the problem is to be considered as not correctly defined (for instance, because of an insufficient definition of constraints).

A mathematical model implemented as an integral equation (2.2) is said to be correctly defined if the model complies with the following characteristics:

1. There exists a unique solution.

2. Small variations of the boundary conditions lead to small variations of the solution.

3. The solution T_i exists for all $\tau_0 \leq t \leq \tau$ and tends to a maximum for $t \to \tau$.

Note, that item 2 is necessary in order to justify that the obtained approximate solution is close to the real one and does not behave chaotically. However, any specific case has to show what is to be understood as "small variation". In addition to the correctness hypotheses, in the present approach it is essential that the validity region $\mathcal{V} := \{(\{c_{i\alpha}\}|t)|\forall c_{i\alpha} \text{ "allowed"} \wedge t \in [\tau_0, t]\} \subset \mathcal{H}$ is a true subset of the domain $\mathcal{H}$ where the correctness definition of the problem holds. The optimum $T_i(\{c_{i\alpha}\}|\tau)$ is supposed to truly lie in the domain and not on the boundary, whereas fragments of the evolution trajectory may cross "forbidden regions" by non-holonomic constraints (i.e. inequalities), which indicates the necessity of a temporary parameter constraint relaxation in order to obtain the desired solution (optimum). Note, that non-holonomic constraints may be useful to restrict the parameter domain right from the beginning, in more complex cases, however, it may be more convenient or even manageable at all to restrict the parameter space in a posterior analysis only (see Figure 1).

6. Conclusions

A self-consistency approach for simulating an abstract dynamical optimization scenario has been discussed based on a parametric inference techniques. The model

has been setup, using pseudo-analytic and local relations which result in a set of first order differential equations. We indicated how to construct the self-consistency function and determine the best estimate parameter set for the problem cast into integral form. Instead of using inference techniques in its original sense, i.e. optimizing trial functions to experimental data, the latter are substituted by a self-consistent equation – the master equation – for the objective function (eq. 2.2). As a main achievement of the present approach we consider the continuous extension of the method to functional samples, which opens up the possibility to apply this kind of approach as a learning procedure for dynamical optimization scenarios, where the relations of parameters, types of local variable dependencies with optimization policies may be explored.

The Self-consistent Parametric Inference algorithm ($\mathcal{SPI}$) is guaranteed to find the solution, if there exists one, and otherwise yields the proof that the problem is insolvable while either based on the proposed parametrization or on the implemented constraints and random number limits. Even after solving the procedure only once, the underlying parametrization may be analyzed for its significance, parameters with almost no influence are indicated as such and may be eliminated in a subsequent post-analysis, which is learning from the previous solution. From the self-consistency measure it is evident, that there is no need for a systematic search through the entire parameter space in order to find the extremal solution, as is usually the case for search algorithms, once the proof of optimality is needed. However, those may be able to find a solution quickly if there are many possible solutions, but according to our experience things get more difficult if there are few solutions, and additionally leave the question open whether a problem has a solution at all or whether the solution is optimal. The present method does not lack from these features, i.e. a genuine proof for the optimum is obtained. For simulation purposes the present procedure has proven to be a valid tool in order to explore models, which may further be classified according to their usefulness or even applicability. We believe, but leave this for a future investigation, that visualizing the parameter hyperspace may especially help modelling scenarios where changes in the original problem have to be taken into account, due to a necessity of problem redefinition.

Although the $\mathcal{SPI}$ method has been tested against Tabu Search implementations in some practical applications, more simulation experience, especially in competition with different algorithms, have to be performed in order to prepare the method as an automated and reliable tool, which may be used not only in simulations but in a variety of practical applications as well. A further topic to be investigated in more detail is the implication of the analyticity of the approach, which by the same formalism permits classical deterministic as well as stochastic descriptions, while performing the neighborhood search. Thus, this feature breaks with the commonly accepted paradigm that analytical and stochastic approaches in one and the same procedure are not compatible.

Resumo. O objetivo deste trabalho é apresentar e discutir um procedimento dinâmico de otimização, baseado no algoritmo de inferência paramétrica autoconsistente, em que os movimentos de uma solução viável para outra são implementados no modo contínuo. Esta abordagem é baseada em uma parametrização genérica do
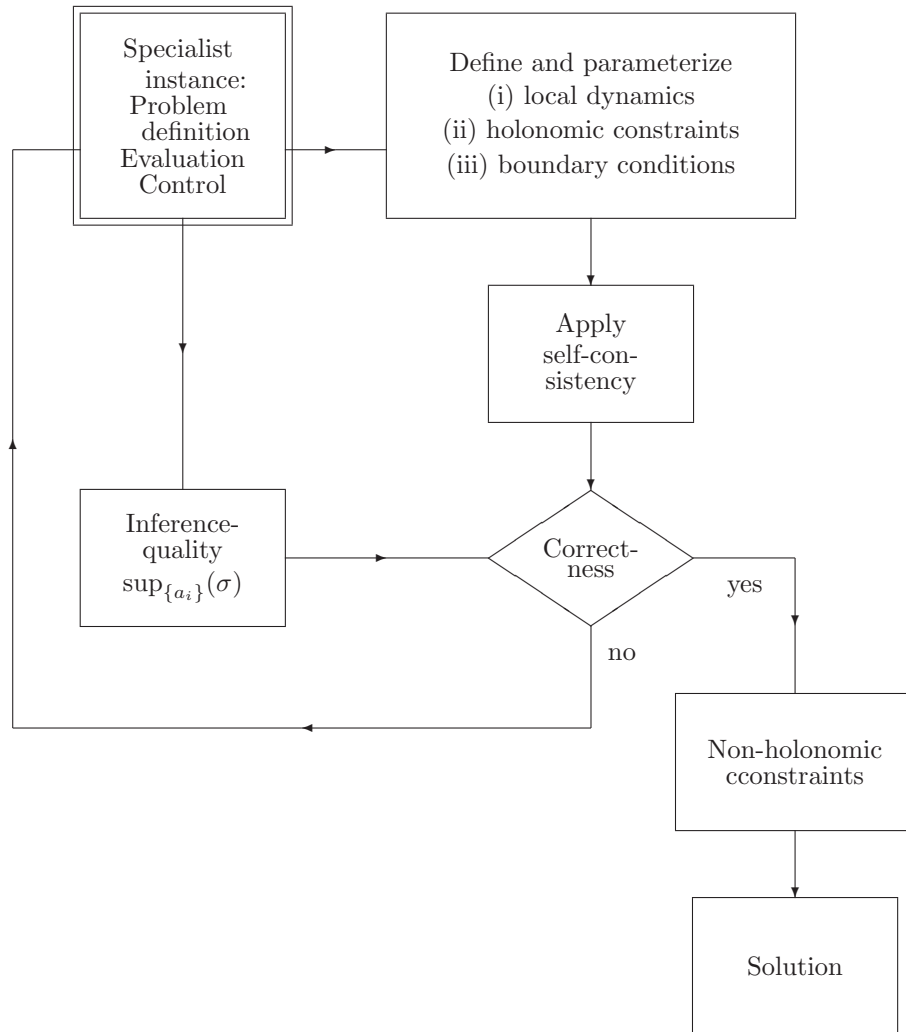
Figure 1: The flow chart diagram of the S̲elfconsistent P̲arametric I̲nference algorithm ($\mathcal{SPI}$).

espaço de pesquisa e utiliza a definição de uma métrica independente de qualquer modelo que guiará o desenvolvimento deste procedimento. Uma das principais contribuições do método aqui proposto é que, encontrada uma solução, esta é ótima dada a parametrização do modelo considerada. Isto é garantido e provado pela autoconsistência.

References

[1] A.T. Gómez, "Modelo para Seqüênciamento de Partes e Ferramentas em um Sistema de Manufatura Flexível com Restrições às Datas de Vencimento e a Capacidade do Magazine", Ph.D. thesis, Instituto Nacional de Pesquisas Espaciais. São José dos Campos, SP, Brazil, 1996.

[2] A.G. Rodrigues, A.T. Gómez, Utilização de metaheurísticas para a modelagem de problemas de seleção de partes e "scheduling" em um sistema de manufatura flexível. *Scientia*, **11** (2000) 17-47.

[3] B.E.J. Bodmann, A.T. Gómez, The self-consistent parametric inference algorithm (SPI): optimization in continuous search space, in "Proceedings of the International Conference on Industrial Engineering and Production Management", IEPM'99 Book 1, pp. 66-77, 1999.

[4] V.J. Rayward-Smith, I.H. Osman, C.R. Reeves and G.D. Smith, "Modern Heuristic Search Methods", Wiley, Chichester, 1996.

[5] G. Roepsdorff, "Path Integral Approach to Quantum Physics", Springer Verlag, Berlin, Heidelberg, New York, 1996.

[6] P.E. Hart, J. Nilsson, B. Raphael, A formal basis for the heuristic determination of minimum cost paths, *IEEE Transactions on SSC*, **4** (1968), 100-107. also *SIGART Newsletter*, **37** (1972) 28-29.

[7] J.C. Furtado, "Algoritmo Genético Construtivo na Otimização de Problemas Combinatórias de Agrupamentos", Ph.D Thesis, pp. 16-34, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP, Brazil, 1998.

[8] M. Tanner, "Tools for Statistical Inference", Springer Verlag, New York, Berlin, Heidelberg, Tokyo, 1996.

[9] G.B. Wetherill, "Intermediate Statistical Methods", pp. 71-73, Chapman and Hall Ltd., London, New York, 1981.

[10] R. Fletcher, "Practical Methods of Optimization", Wiley & Sons, New York, 1987.

[11] J. Kohlas, "Zuverlässigkeit und Verfügbarkeit: Mathematische Modelle, Methoden und Algorithmen", Teubner Verlag, Stuttgart, 1987.

[12] J. Guddart, F. Guerra Vasquez, H. Jongen, "Parametric Optimization: Singularities, Pathfollowing and Jumps", Teubner Verlag, Stuttgart, 1990.

[13] E. Coffman, J. Lenstra, A. Rinnooy Kan, "Handbooks in Operation Research and Management Science", Elsevier, Amsterdam, 1992.

[14] L. Fould, "Graph Theory and Applications", Springer Verlag, New York, 1992.

[15] J. Aubin, "Optima and Equilibria", Springer Verlag, Berlin, 1993.

[16] M. Grötschel, L. Lovász, A. Schrijver, "Geometric Algorithms and Combinatorial Optimization", Springer Verlag, New York, 1993.

[17] J. Hiriart-Urruty, C. Lemarchal, "Convex Analysis and Minimization Algorithms", Springer Verlag, New York, 1993).

[18] E. Kreyszig, "Introductory Functional Analysis with Applications", John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore, 1978.

[19] J.B. Marion, S.T. Thornton, "Classical Dynamics of Particles and Systems", pp. 225-227, Saunders College Publishing, San Diego, New York, 1995.