Optimization with State Space Approaches II: System Theoretic Concepts per Example

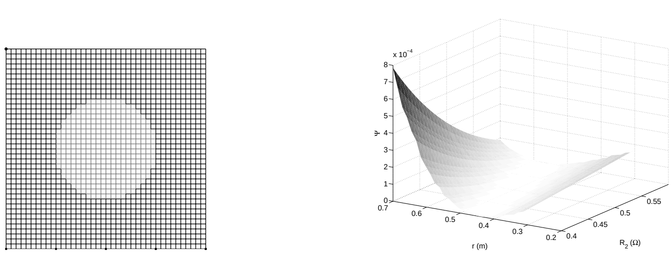
M. Neumayer and G. Steiner
Institute of Electrical Measurement and Measurement Signal Processing
Kopernikusgasse 24/IV, A-8010 Graz, Graz University of Technology, Austria
Email: neumayer@tugraz.at

***Abstract*—The development and hence the number of deterministic, stochastic and/or hybrid optimization algorithms has seen a rigorous grow in the past years and is still of major concern in the optimization community. In system and control theory engineers use state space models as an unified basis for analysis and design tasks. In this paper the advantageous use of state space techniques for optimization is demonstrated on two examples. In particular the handling of models errors due to reduced models and a hybrid optimization algorithm are presented.**

I. Introduction

In the past years an enormous variety of different deterministic, stochastic or hybrid labeled optimization algorithms has grown up in the optimization community, making a distinction between the different concepts more difficult and, more important, the selection of a suitable and efficient algorithm often hard. This work builds the practical extension to its companion paper [1], where the use of state space models and methods for optimization is proposed. State space models provide a unified representation of dynamic systems. Considering optimization algorithms as dynamic systems, the approach provides the facility of a framework for comparative analysis and the design of new algorithms. Further, advantageous techniques of state estimation and control can be employed. In this paper, we demonstrate the use of state space techniques for optimization by two examples.

II. A Resistor Network Problem



(a) Resistor net. (b) Objective function Ψ.

Fig. 1. Test example and solutions.

Figure 1(a) depicts a resistor network, which will be the basis for all demonstrations. The black lines represent resistors of $1\,\Omega$. The gray resistors lines form a circular object of radius r and have the resistor value R_2. It is aim to determine the radius r and the resistor value R_2 by boundary measurements. As the mapping of the circle is done using the center coordinates of the resistors, the problem is discontinuous for the variable r. To take boundary measurements, a voltage source is connected to the upper left corner of the network and several current measurements are taken on some points on the lower edge, take obtain the current distribution. Hence, for the state vector $\boldsymbol{x} = \left[\begin{array}{cc} r & R_2 \end{array} \right]^T$ the measurement vector is given by $\boldsymbol{y} = \left[\begin{array}{ccc} I_1 & \dots & I_N \end{array} \right]^T$. Figure 1(b) depicts the typical shape of the two dimensional objective function

$$\Psi(\boldsymbol{x}) = (\boldsymbol{y} - \boldsymbol{y}_{true})^T(\boldsymbol{y} - \boldsymbol{y}_{true}), \qquad (1)$$

in the region near the solution $\boldsymbol{x}_{true}$. $\boldsymbol{y}_{true}$ collects the currents of the true model. In our example the parameters were set to $R_2 = 0.5\,\Omega$ and $r = 0.5\,\mathrm{m}$. The corner length of the network was $2\,\mathrm{m}$.

III. Two Examples

A. Handling Errors of Reduced Models

Recently, in order the decrease the computational load due to complex forward models, the use of approximated or reduced forward models has become more and more popular in optimization [2]. E.g. a simple reduction technique is given by the use of a coarser finite element discretization. However, this has also an effect on the objective function $\Psi(\boldsymbol{x})$, which is now replaced by the objective function $\Psi^*(\boldsymbol{x})$ due to the reduced model. As the reduced model may have different characteristics compared to reality, the reduction has to be handled carefully. Similar to the so-called enhanced error model [3], the following equation can be established

$$\Psi^*(\boldsymbol{x}) = \Psi(\boldsymbol{x}) + (\Psi^*(\boldsymbol{x}) - \Psi(\boldsymbol{x})) = \Psi(\boldsymbol{x}) + v(\boldsymbol{x}). \qquad (2)$$

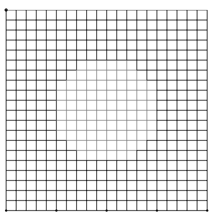
Hence, the model error $v(\boldsymbol{x})$ can be understood as an additive error term. Although $v(\boldsymbol{x})$ is a strictly deterministic term, it is common to treat such errors like random variables and approximate them by means of a Gaussian distribution [3] using a mean μ_v and a covariance $\boldsymbol{\Sigma}_v$. In general their determination would require a sampling procedure. However, during the setup period of the model mostly information about its error (or uncertainty) comes to hand, which can be used to characterize $v(\boldsymbol{x})$. Now, Bayes' law

$$\pi(\boldsymbol{x}|\boldsymbol{y}) = \frac{\pi(\boldsymbol{y}|\boldsymbol{x})\pi(\boldsymbol{x})}{\pi(\boldsymbol{y})} \propto \pi(\boldsymbol{y}|\boldsymbol{x})\pi(\boldsymbol{x}), \qquad (3)$$

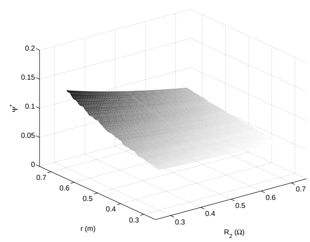
can be applied to compute the posteriori probability $\pi(\boldsymbol{x}|\boldsymbol{y})$. As $\pi(\boldsymbol{x})$ can be understood as a function which holds the constraints, the likelihood function $\pi(\boldsymbol{y}|\boldsymbol{x})$ is given by

$$\pi(\boldsymbol{y}|\boldsymbol{x}) \propto \exp\left(-(\Psi^* - \mu_v)^T \boldsymbol{\Sigma}_v^{-1} (\Psi^* - \mu_v)\right) \qquad (4)$$

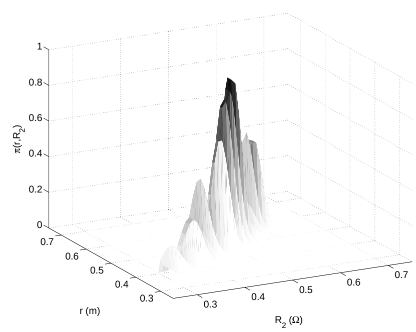
builds an equivalent to the objective function. Figure 2(a) depicts a coarse resistor network which is used for estimating $\boldsymbol{x}$ out of measurements of the fine resistor network depicted in Figure 1(a). The objective function Ψ^* is depicted in Figure 2(b). The reduction is already too rough and it would be impossible to obtain the right result out of Ψ^*. By applying the enhanced error approach in combination with Bayes law, one can obtain Figure 2(c), where the probability $\pi(r, R_2)$ is depicted. Out of this, the solution can be found, which matches the true values. It will practically not happen, that the reduction



(a) Reduced resistor net.



(b) Objective function Ψ^*.



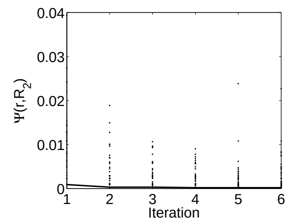
(c) Posteriori probability $\pi(r, R_2)$.

Fig. 2. Determination of r and R_2 using a reduced model.

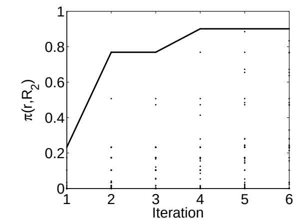
is as rough like the one in the example. However, the example illustrates the increased possibility of the approach.

B. A Hybrid Optimization Scheme.

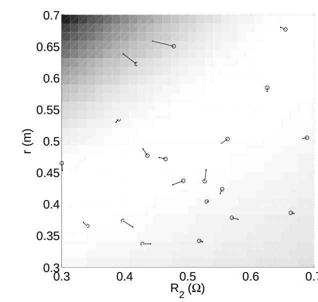
As a second example we want to present a hybrid version of the steepest descent algorithm. Again, we want use the reduced model depicted in figure 2(a). As explained in [1], deterministic optimization algorithms have similarities to feedback loops. For the steepest descent algorithm, the control variables are given by $\boldsymbol{u}_k = -s\boldsymbol{g}(\boldsymbol{x}_k)$, where $\boldsymbol{g}(\boldsymbol{x}_k)$ denotes the gradient of the objective function and s is a step width. As Ψ is discontinuous for the variable r, we will now treat its component of the gradient as random variable. Hence, a hybrid optimization algorithm can be obtained. For the optimization, we use a particle filter [4]. The principle of this state estimator is similar to optimization methods like Particle Swarm Optimization, or differential evolution. In a first step several proposal candidates are generated. After evaluating their fitness a resampling procedure is performed. Then the procedure is repeated for the resampled candidates. Figure 3(c) depicts particles of the particle filter at an initial time step exploring the function space. Figure 3(a) and 3(b) depict the trend of the objective function and the posteriori distribution. The particle filter is able to find the right solution after few iterations.



(a) Objective function.



(b) Posteriori probability.



(c) Particles exploring the function.

Fig. 3. Output of the particle filter.

IV. Conclusion

In this paper two examples of using state space methods and concepts for optimization and their potential are presented. Example one focused on the ability to use reduced models in combination with an error model. Example two presented a hybrid optimization concept, where the steepest descent algorithm was implemented as state space system. The final paper will present more details about the implementations, as well as some further details concerning the application of state space techniques for optimization.

References

[1] M. Neumayer and G. Steiner, *Optimization with State Space Approaches I: A System Theoretic View on Optimization Algorithms*, submitted to 18^{th} International Conference on the Computation of Electromagnetic Fields, COMPUMAG 2011.

[2] M.N. Albunni, V. Rischmuller, T. Fritzsche, and B. Lohmann, *Multiobjective Optimization of the Design of Nonlinear Electromagnetic Systems Using Parametric Reduced Order Models*, IEEE Transactions on Magnetics, vol.45, no.3, pp.1474-1477, March 2009.

[3] J. P. Kaipio and E. Somersalo, *Statistical and computational inverse problems*, New York: Applied Mathematical Sciences, Springer, 2004.

[4] M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, *A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking* IEEE Transactions on Signal Processing 50 (2), pp.174188, 2002.