Two new variants of the manifold-mapping technique

D. Echeverria

Centrum voor Wiskunde en Informatica (CWI), Amsterdam, The Netherlands

Abstract

Purpose – Manifold-mapping (MM) is an efficient surrogate-based optimization technique aimed at the acceleration of very time-consuming design problems. In this paper we present two new variants of the original algorithm that make it applicable to a broader range of optimization scenarios.

Design/methodology/approach – The first variant is useful when the optimization constraints are expressed by means of functions that are very expensive to compute. The second variant endows the original scheme with a trust-region strategy and the result is a much more robust algorithm.

Findings – Two practical optimization problems from electromagnetics eventually show that the proposed variants perform efficiently.

Originality/value – The original MM algorithm is extended with two new variants. Therefore, the MM approach is applicable to a much larger set of design situations.

Keywords Optimization techniques, Electromagnetic fields

Paper type Research paper

1. Introduction

Space mapping (Bandler et al., 1994, 2004; Echeverria et al., 2005) was introduced by J.W. Bandler as a surrogate-based optimization technique aimed at optimization processes where very accurate (and quite often time-consuming) models are needed. In space-mapping terminology these models are called fine models. The surrogate is a correction of a fast-to-compute approximation of the fine model. This correction and the approximation are commonly known as the space-mapping function and as the coarse model, respectively.

However, the space-mapping approach does not in general yield the correct optimum. Moreover, the computation of the space-mapping function requires an additional optimization problem. This minimization process is sometimes problematic (Bandler et al., 2004). By manifold-mapping (MM) (Echeverria and Hemker, 2005), we overcome those difficulties and we obtain an improved and efficient optimization strategy.

The MM concept introduces an alternative correction of the coarse model. Unlike the space-mapping function, it is very straightforward to compute and the associated MM algorithm can be proven to converge to the true optimum (Echeverria and Hemker, 2006). The MM algorithm has been reported as an efficient optimization technique in problems of practical relevance (Echeverria et al., 2006).

The author would like to thank P.W. Hemker for his kind and detailed feedback to this manuscript, J. ter Maten, C. Lin and J. Ramos for their help concerning the class-E power amplifier problem and J. Werner for her valuable proofreading. This research is supported by the Dutch Ministry of Economical Affairs within the project IOP-EMVT 02201.
In this paper, we present two significant improvements of the original MM scheme. The first one allows its use in design scenarios where optimization constraints are expressed by functions that are very expensive to compute. The second improvement endows the original iteration with a trust-region strategy that increases the algorithm robustness, when for example, the optimization problem is strongly nonlinear or ill-conditioned (i.e. multiple optima are present).

This paper is structured as follows. In Section 2, the basics of MM are explained. The two new variants, constrained optimization with MM and trust-region manifold-mapping (TRMM), are introduced in Section 3. Eventually in Section 4, these two new algorithms are applied to practical design problems from electromagnetics.

2. Manifold-mapping basics
Before explaining the basics of the MM approach, we introduce the notation adopted. The design specifications are denoted by \( y \in \mathbb{R}^m \). The fine model response \( f(x) \) is defined over the set \( X \subseteq \mathbb{R}^n \) and \( x \in X \) is the design variable. In this work, we minimize the discrepancy between the specifications and the fine model response:

\[
x_f^* = \arg \min_{x \in X} \| f(x) - y \|.
\]

We refer to \( x_f^* \) as the fine model optimum. Not every optimization problem is of this type, but most practical design situations can be reduced to this structure.

The coarse model response is denoted by \( c(x) \) and it is also defined over \( X \). We can consider coarse models over sets \( Z - X \) by the introduction of an additional mapping \( \hat{p} : X \rightarrow Z \), as in Echeverria and Hemker (2005). The coarse model optimum is defined as:

\[
x_c^* = \arg \min_{x \in X} \| c(x) - y \|.
\]

A general constrained optimization problem can be stated with the help of a constraint function \( k_f(x) \) as:

\[
x_f^* = \arg \min_{x \in X} \| f(x) - y \|, \quad \text{subject to } k_f(x) \geq 0.
\]

Now \( f \) and \( k_f \) are defined over \( \hat{X} \subseteq \mathbb{R}^n \), a superset of \( X \). The function \( k_f \) refers to the fine (accurate) equality and inequality constraints. In this section \( k_f \) will be assumed easy-to-compute and thus, no special care with respect to the constraints has to be taken in the MM approach. For those cases where \( k_f \) requires a significant computational effort, MM can still be applied whenever a fast-to-evaluate function \( k_c \) is available for approximating the fine constraint function \( k_f \) (Section 3.1).

In Echeverria and Hemker (2005) the manifold mapping \( S : c(X) \rightarrow f(X) \) is introduced with the aim of correcting the misalignment between the models \( f \) and \( c \). With \( S \), the point \( c(x_f^*) \) is mapped to \( f(x_f^*) \) and the tangent plane for \( c(\hat{X}) \) at \( c(x_f^*) \) to
the tangent plane for \( f(X) \) at \( f(x^*_f) \) (Figure 1). Other approaches are possible but in our work we define \( S \) as the affine mapping:

\[
S \ c(x) = f(x^*_f) + \bar{S}(c(x) - c(x^*_f)),
\]

where:

\[
\bar{S} = f'_f(x^*_f)J^{-\dagger}_c(x^*_f).
\]

The \( \dagger \) symbol represents the pseudoinverse and \( J_f(x^*_f) \) and \( J_c(x^*_f) \) denote, respectively, the Jacobian of \( f \) and \( c \) at \( x^*_f \). The combination \( S \circ c \) acts as the surrogate model for \( f \). Under some assumptions about the similarity between the fine and the coarse model that usually hold in practice, it can be shown (Echeverría and Hemker, 2006) that

\[
x^*_f = \arg \min_{x \in X} \|S \ c(x) - y\|
\]

\[
= \arg \min_{x \in X} \|c(x) - c(x^*_f) + \bar{S}^\dagger \left(f(x^*_f) - y\right)\|.
\]

The last equality is the basis for the MM algorithm.

The mapping \( S \) is not known a priori, because it depends on the solution of the optimization problem. But it can be iteratively approximated. The MM algorithm (Figure 2) yields both the mapping \( S \) and the desired fine model optimum \( x^*_f \) (Echeverría and Hemker, 2006). This scheme is particularly attractive from an implementational point of view: the central function in every minimization process is always the coarse model \( c(x) \) and not a surrogate (the specifications are updated in each step).

3. Two new manifold-mapping variants

The MM algorithm (Figure 2), in the previous section, cannot be applied efficiently when the constraint function \( k_f \) is expensive to compute. Moreover, it can fail when the models are ill-conditioned, i.e. a small variation of the design variable does not affect

---

Figure 1. Manifold-mapping model alignment
the model response. Therefore, we present two new variants of the basic scheme in order to cope with these two problematic situations.

3.1 Constrained optimization with manifold mapping

If a fast-to-evaluate approximation \( k_c \) of the fine constraint function \( k_f \) is available, the MM idea can be extended to a constrained optimization problem in a straightforward way. An additional (affine) mapping \( K : k_c(\tilde{X}) \rightarrow k_f(\tilde{X}) \) between the constraint sets, \( k_c(\tilde{X}) \) and \( k_f(\tilde{X}) \) is introduced. With \( K \), the point \( k_c(\tilde{x}_f^*) \) is mapped to \( k_f(\tilde{x}_f^*) \) and the tangent manifold for \( k_c(\tilde{X}) \) at \( k_c(\tilde{x}_f^*) \) to the tangent manifold for \( k_f(\tilde{X}) \) at \( k_f(\tilde{x}_f^*) \). Again, the mapping \( K \) is not known a priori and therefore, it has to be approximated iteratively. We indicate in Figure 3, how the MM algorithm has to be modified in order to deal with constraints that are expensive to evaluate (the approximated constraint manifold mapping is denoted by \( K_k \)). The stationary point of the iteration is the right optimum \( x_f^* \) since in the limit, the Karush-Kuhn-Tucker conditions (Nocedal and Wright, 1999) hold.

3.2 Trust-region manifold mapping

An important requirement for convergence of the MM algorithm is that the matrices \( \Delta F \) and \( \Delta C \) are sufficiently well-conditioned (Echeverria and Hemker, 2006). In practice, it is hard to foresee the effect of this requirement. In the initial phase of the
whole optimization process we cannot give any a priori guarantee concerning the condition number associated to the mentioned matrices.

When these matrices are ill-conditioned, the algorithm can take very large steps because of the pseudo-inverse computation. Therefore, in Hemker and Echeverria (2006), a trust-region strategy (Conn et al., 2000) is introduced in such a way that the undesired effects do no harm and convergence is accelerated. The matrices $\Delta F$ and $\Delta C$ are regularized with the help of a generalized singular value decomposition (GSVD) of the pair $(\Delta F, \Delta C)$.

Figure 3. The manifold-mapping algorithm for constrained optimization.

```plaintext
x_0 = x^*_c = \arg\min_{x \in \mathcal{X}} \|c(x) - y\| \text{ subject to } k_c(x) \geq 0;
T_0 = I_{m \times m};
K_0 \cdot = k_f(x_0) + (\cdot - k_c(x_0));
for \ k = 0, 1, \ldots, \ while \ ldots
\begin{align*}
  y_k &= c(x_k) - T_k (f(x_k) - y); \\
  x_{k+1} &= \arg\min_{x \in \mathcal{X}} \|c(x) - y_k\| \text{ subject to } K_k(k_c(x)) \geq 0;
\end{align*}
break if \ldots
\begin{align*}
  \Delta F &= [f(x_{k+1}) - f(x_k), \ldots, f(x_{k+1}) - f(x_{\max(k+1-n,0)})]; \\
  \Delta C &= [c(x_{k+1}) - c(x_k), \ldots, c(x_{k+1}) - c(x_{\max(k+1-n,0)})]; \\
  \Delta K_f &= [k_f(x_{k+1}) - k_f(x_k), \ldots, k_f(x_{k+1}) - k_f(x_{\max(k+1-n,0)})]; \\
  \Delta K_c &= [k_c(x_{k+1}) - k_c(x_k), \ldots, k_c(x_{k+1}) - k_c(x_{\max(k+1-n,0)})]; \\
  (U_{\Delta F}, \Sigma_{\Delta F}, V_{\Delta F}) &= \text{SVD}(\Delta F); \\
  (U_{\Delta K_c}, \Sigma_{\Delta K_c}, V_{\Delta K_c}) &= \text{SVD}(\Delta K_c); \\
  \Delta F^\dagger &= V_{\Delta F} \Sigma_{\Delta F}^1 U_{\Delta F}^T; \\
  \Delta K_c^\dagger &= V_{\Delta K_c} \Sigma_{\Delta K_c}^1 U_{\Delta K_c}^T; \\
  T_{k+1} &= \Delta C \Delta F^\dagger; \\
  K_{k+1} \cdot &= k_f(x_{k+1}) + \Delta K_f \Delta K_c^\dagger (\cdot - k_c(x_{k+1}));
\end{align*}
enddo
```
The GSVD of the pair of matrices \((\Delta F, \Delta C)\) is (Golub and van Loan, 1983) a set of five matrices \(U_{\Delta F}, U_{\Delta C}, \Sigma_{\Delta F}, \Sigma_{\Delta C}\) and \(V\), such that

\[ \Delta F = U_{\Delta F} \Sigma_{\Delta F} V^T \quad \text{and} \quad \Delta C = U_{\Delta C} \Sigma_{\Delta C} V^T, \]

with \(V\) a regular matrix, \(U_{\Delta F}\) and \(U_{\Delta C}\) unitary matrices and \(\Sigma_{\Delta F} = \text{diag}(\sigma_{1}^{\Delta F}, \ldots, \sigma_{n}^{\Delta F})\) and \(\Sigma_{\Delta C} = \text{diag}(\sigma_{1}^{\Delta C}, \ldots, \sigma_{n}^{\Delta C})\) diagonal matrices with non-negative elements. The matrix \(\Delta C \Delta F^\dagger\) in the MM algorithm can now be written as

\[ \Delta C \Delta F^\dagger = U_{\Delta C} \Sigma_{\Delta C} V^T (V^T)^{-1} \Sigma_{\Delta F}^+ U_{\Delta F}^T = U_{\Delta C} \text{diag} \left( \frac{\sigma_{i}^{\Delta C}}{\sigma_{i}^{\Delta F}} \right) U_{\Delta F}^T, \]

with \(\lambda \geq 0\).

Regularization is introduced by taking \(\lambda > 0\). The choice of the parameter \(\lambda\) in the TRMM algorithm (Figure 4) is based on the success of the previous iteration steps. If the residual \(\|f(\mathbf{s}_{k+1}) - \mathbf{y}\|\) decreases, the value of \(\lambda\) is divided by \(r_\lambda\), otherwise it is multiplied by \(R_\lambda\). In practice, \(\lambda\) is never reduced below a tolerance \(\tau\). The values suggested for a few constants in the algorithm are based on heuristics. Further, the scheme contains an (optional) damping parameter \(\delta \geq 0\). For strongly nonlinear problems it can stabilize the convergence process at the expense of a few additional function evaluations (Hemker and Echeverría, 2006).

4. Optimization experiments

The two design problems in this section have practical relevance. The first problem has an expensive constraint function and thus, it is suitable for the MM variant presented in Section 3.1. The second one is the well-known TEAM problem 25 for which we will see that the trust-region strategy in Section 3.2 makes a significant improvement on the original MM algorithm.

Two optimization methods are used for the function \(\arg \min\) in the schemes introduced in this paper: sequential quadratic programming (SQP) (Nocedal and Wright, 1999) and differential evolution (DE) (Storn and Price, 1995). We use SQP and DE for local and global minimization purposes, respectively.

4.1 A class-E power amplifier

In many electronic circuits, the proper device operation is specified by a number of constraints that very often are expensive to evaluate (for example, when the voltage in a node should be larger than a certain threshold).

Class-E power amplifiers (Sokal and Sokal, 1975; Ramos, 2005) are widely used in the radio frequency and microwave field because of their high efficiency. In the amplifier in Figure 5 we see a transistor \(M_1\) followed by a resonant filter \(L_0 - C_0\). The transistor acts as a switch and drives the current through the inductor \(L_1\).

The design problem consists in the minimization of the supplied power (i.e. the model response is a scalar and \(\mathbf{y} = 0\)). The constraints are defined over \(\mathbb{R}^3\) and they ensure that the circuit is an amplifier that delivers an output power larger than a minimum value. We need to compute the current through \(L_1\) in order to obtain the supplied power (the supplied voltage \(V_{DD}\) is constant and equal to 2 V). The constraints
are based on the voltages in nodes P0 and P1. The nature of the amplifier implies in all cases expensive transient simulations. The design variable $x_1$ is the internal drawn channel width of the transistor, $x_2$ and $x_3$ the inductances $L_1$ and $L_0$, respectively, $x_4$ the capacitance $C_0$ and $x_5$ the load resistance $R_L$. For a more complete problem description see Ramos (2005).

The fine model response and constraint function, $f$ and $k_f$, result from a transient simulation where the input wave is a square signal of 50 cycles with a time resolution of 1,000 samples per cycle[1]. The cost function is $F(x) = |f(x) - y|$. The coarse model
and constraint function, \( c \) and \( k_c \), use only 5 cycles with 100 samples per cycle. In both cases, the backward Euler scheme is used for the integration of the corresponding differential equation. In the second situation the computation of the model and the constraint function is 100 times faster compared to the first one. In Figure 6, we show the transient voltages (last cycle) in node \( P_1 \) for the fine and the coarse simulations with \( x = x^*_f \). Though the differences may seem small, \( x^*_f \) is only feasible for the fine constraints.

The MM variant for constrained optimization yields the solution \( x = [17,530 \ \mu m, 3.54 \ \text{nH}, 1.33 \ \text{nH}, 63.1 \ \text{pF}, 3.03 \ \text{V}] \) after 26 equivalent fine transient simulations. (This value is approximately proportional to the total computing time.) The associated (fine) supplied power is 0.694 \( \text{W} \). We use a combination of DE and SQP for the coarse
model optimum computation and only SQP for the rest of the minimization processes within MM. The solution given in Ramos (2005) is also feasible and shows a supplied power of 0.774 W. But it is computed with DE and therefore, a few thousand simulations are needed. SQP, with the coarse model optimum as initial guess, finds a feasible point (associated supplied power of 0.700 W) after 56 equivalent fine transient simulations.

Subsequent optimization runs detect different design solutions with analogous quality. The costs to compute them are also similar to those indicated above. The presence of multiple optima is a clear sign of ill-conditioned model responses and/or constraint functions. The influence of this fact in the performance of the MM scheme has been almost imperceptible. More difficult cases could be dealt with by means of a proper trust-region strategy.

4.2 A die press (TEAM problem 25)

We will study the optimization of a die press with an electromagnet (TEAM Workshop problem 25) (Takahashi et al., 1996). This device (Figure 7) is used for manufacturing anisotropic permanent magnets. The molds and the pole are made of steel.

The specifications \( y \) are the magnetic flux computed in ten points along the curve \( e-f \) (Figure 7) in the cavity; i.e. \( y = [B_x; B_y] \) where \( B_x \) and \( B_y \) are the following column vectors:

\[
B_x = [0.35 \cos \theta_i]_{i=1,...,10} \quad \text{(T)},
\]

\[
B_y = [0.35 \sin \theta_i]_{i=1,...,10} \quad \text{(T)},
\]

and \( \theta_i \) denotes the angle with the \( x \)-axis. The design variable \( x = [x_1, x_2, x_3, x_4] \) refers to the geometry of the molds. The control space \( X \) is a polytope in \( \mathbb{R}^4 \) (inexpensive constraint function). The complete problem description can be found in Takahashi et al. (1996).

The fine model \( f \) is based on second order triangular finite elements with four levels of adaptive refinement (this yields around 120,000 degrees of freedom). The cost function is \( F(x) = ||f(x) - y||_2^2 \). The coarse model \( c \) is built as a least squares quadratic approximation of 16 finite element solutions with a linearized characteristic of the molds and the pole. The 16 points in the approximation are vertices of the polytope \( X \). The discretizations used there are based on very coarse grids (with less than 1,000 degrees of freedom). The construction and evaluation of the coarse model has a negligible computational cost when compared with that of the fine model.

Table I shows the die press optimization results. We observe that MM after the fourth iteration seems to move away from the optimum. This diverging tendency grows during the next iterations. The solution \( x = [7.14, 13.70, 14.00, 14.33] \) (mm) is computed by TRMM in five iterations (eight equivalent fine model evaluations). The associated cost function value is 0.0004 and the quality of the solution is acceptable for practical purposes. We use DE for the coarse model optimum computation and for every minimization process within the MM approach. TRMM performs much more efficient than SQP, which was applied with the coarse model optimum as initial guess.
5. Conclusions

MM is a multi-level surrogate-based optimization strategy aimed at very time-consuming design problems. In this paper, we have presented two new variants of the original MM algorithm. The first one is useful when the optimization constraints are given by functions that are very expensive to compute. The second one incorporates a trust-region approach that yields a much more robust scheme. Both variants perform efficiently in solving practical design problems from electromagnetics.

Note
1. The frequency of operation is 850 MHz.

References


**Corresponding author**

D. Echeverría received the Telecommunication Engineering degree from the University of Zaragoza (Spain) in 1998. In 2002, he obtained the MSc degree in Applied Mathematics from the same university. Since, April 2003, he is affiliated with the Centre for Mathematics and Computer Science (CWI) in Amsterdam (The Netherlands). D. Echeverría can be contacted at: D.Echeverria@cwi.nl

---

To purchase reprints of this article please e-mail: reprints@emeraldinsight.com
Or visit our web site for further details: www.emeraldinsight.com/reprints