# OPTIMA 79

Mathematical Programming Society Newsletter


Katya Scheinberg

## Geometry in model-based algorithms for derivative-free unconstrained optimization

*Abstract.* Derivative free optimization addresses general nonlinear optimization problems in the cases when obtaining derivative information for the objective and/or the constraint functions is impractical due to computational cost or numerical inaccuracies. Applications of derivative free optimization arise often in engineering design such as circuit tuning, aircraft configuration, water pipe calibration, oil reservoir modeling, etc. Traditional approaches to derivative free optimization until late 1990's have been based on sampling of the objective function, without any attempt to build models of the function or its derivatives. In the late 90's model based trust region derivative free methods started to gain popularity, pioneered by Powell and further advanced by Conn, Scheinberg and Toint. These methods build linear or quadratic interpolation model of the objective function and hence can exploit some first and second oder information. In the last several years the general convergence theory for these methods, under reasonable assumptions, was developed by Conn, Scheinberg and Vicente. Moreover, recently Scheinberg and Toint have discovered the "self-correcting" property which helps explain the good performance observed in these methods and have shown convergence under very mild requirements.

### 1 What is derivative free optimization?

Derivative free optimization is a class of nonlinear optimization methods which usually comes to mind when one needs to apply optimization to complex systems. The complexity of those systems manifests itself in the lack of derivative information (exact or approximate) of the functions under consideration. What usually causes the lack of derivative information is the fact that the function values are a result of a black-box simulation process or a physical experiment. The situation is often aggravated by the high cost of the function evaluations and the numerical noise in the resulting values. Thus the use of finite difference derivative approximation is typically prohibitive.

The numerous applications of derivative free optimization can be found in engineering design, geological modeling, finance, manufacturing, biomedical applications and many other fields. As the available computational power grows the simulation processes become routine and using optimization of complex systems becomes possible and desirable. Thus the number of applications of derivative free optimization grows continuously, which partially explains the continuing growth of the field itself. Another reason for the growth of the field is the recent development of relatively sophisticated algorithms and theory which address the specific needs of the derivative free problems. Here we will discuss some of the recent developments in the theory of model based derivative free methods. We would like to note that the purpose of this article is to focus on the issue of the maintenance of the geometry of the sample sets in model based derivative free methods. Since this is not a survey the list of references is very limited.

### 2 The role of geometry

When it comes to the derivative free optimization, it is clear that most standard optimization approaches do not apply since they rely on Taylor type models and hence require derivatives. Instead various methods of sampling the objective function have been proposed. The most widely used and well-known of them is the Nelder-Mead algorithm [12], popular for its simplicity and effectiveness, but at the same time notorious for its failure to converge even in simple cases.

Roughly speaking, what Nelder-Mead does is the following (for description and analysis of the method see [9] and [23]): the objective function is evaluated at $n+1$ affinely independent points and the point with the worst function value is selected. The worst point is then reflected with respect to the hyperplane formed by the remaining $n$ points. Depending on the function value achieved at this new sample point the original simplex may be contracted or new simplex may be expanded or contracted along a certain direction and a possible new sample point may be produced and evaluated. The contraction and reflection steps are designed to find progress along a (hopefully) descent direction. In the process the shape of the simplex changes, often adapting itself to the curvature of the objective function. This observed behavior of Nelder-Mead method is what makes it often so successful in practice. However, it is also the cause of its failure to converge (in theory and in practice) – the simplex may change the shape until it becomes "flat" and the further progress is impossible because the sample points are no longer affinely independent and the sample space may become orthogonal to the gradient direction.

**Contents of Issue 79 / May 2009**

In the early 90s a new class of derivative free methods emerged – the pattern search methods, suggested by Torczon ([21], [22]). As opposed to the Nelder-Mead method the pattern search methods evaluate the objective function on a pattern of a fixed shape. The pattern can contract or expand, but the shape and, hence, the affine independence of the sample points never change. With this restriction comes the benefit of a global convergence theory, but also the loss of the ability to use the curvature information. The pattern search methods and also a related class of direct search methods, [1], have since grew in variety and sophistication allowing a use of different patterns and incorporating other sampling techniques including the use of interpolation models. However the convergence theory essentially relies on the predetermined geometry of a pattern.

In the mid to late 90s some of the more classical methods from the derivatives based literature found their analogue in the derivative free world. Specifically, pioneered by Powell ([13], [14], [15], [17], [18]), and also developed by Conn, Scheinberg and Toint [10], [4] [5], a class of trust-region methods based on quadratic interpolation, rather than Taylor models, was introduced. Quadratic interpolation models are built based on sample sets of points, preferably in reasonably close proximity to the current best iterate. Due to the expense of the function evaluations, the sample sets typically consist of past iterates, recent unsuccessful steps and possibly some additional sample points. It was understood early on that, unlike the Taylor model, whose accuracy depends entirely on the properties of the approximated function and the distance to the center of the Taylor expansion, the interpolation model's quality depends also on the geometry of the sample set. It was also understood that, if no special care is taken, the sample set may deteriorate, just as in the Nelder-Mead algorithm, and produce incorrect or inaccurate models. It was, thus, believed that the geometry of the sample set needs to be maintained throughout the progress of the algorithm by the means of special "geometry steps".

We will now explain the effect of the geometry on the convergence properties of model-based derivative free methods.

## 3 Interpolation models and trust-region methods

We consider the unconstrained minimization problem

$$\min_{x\in\mathbf{R}^n} f(x) \tag{3.1}$$

where the first derivatives of the objective function $f(x)$ are assumed to exist and be Lipschitz continuous. However, explicit evaluation of these derivatives is assumed to be impossible, either because they are unavailable or because they are too costly.

### 3.1 Polynomial interpolation and Lagrange polynomials

Let us consider $\mathcal{P}_n^d$, the space of polynomials of degree $\leq d$ in $\mathbf{R}^n$ and let $p_1 = p+1$ be the dimension of this space. One knows that for $d=1$, $p_1 = n+1$ and that for $d=2$, $p_1 = \frac{1}{2}(n+1)(n+2)$. A basis $\Phi = \{\phi_0(x), \phi_1(x), \ldots, \phi_p(x)\}$ of $\mathcal{P}_n^d$ is a set of $p_1$ polynomials of degree $\leq d$ that span $\mathcal{P}_n^d$. For any such basis $\Phi$, any polynomial $m(x) \in \mathcal{P}_n^d$ can be written as

$$m(x) = \sum_{j=0}^{p} \alpha_j \phi_j(x),$$

where the $\alpha_j$'s are real coefficients. We say that the polynomial $m(x)$ interpolates the function $f(x)$ at a given point $y$ if $m(y) = f(y)$.

Assume now we are given a set $\mathcal{Y} = \{y_0, y_1, \ldots, y_p\} \subset \mathbf{R}^n$ of interpolation points, and let $m(x)$ denote a polynomial of degree $d$ in $\mathbf{R}^n$ that interpolates a given function $f(x)$ at the points in $\mathcal{Y}$. The interpolation polynomial exists and is unique if and only if the set $\mathcal{Y}$ is *poised*. If the set is poised then one can define the basis of Lagrange polynomials ([16]).

**Definition 3.1.** *Given a poised set of interpolation points $\mathcal{Y} = \{y_0, y_1, \ldots, y_p\}$, a basis of $p_1 = p+1$ polynomials $\ell_j(x)$, $j = 0, \ldots, p$, in $\mathcal{P}_n^d$, is called a basis of Lagrange polynomials if*

$$\ell_j(y_i) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Lagrange polynomials have a number of useful properties. In particular, we are interested in the crucial fact that, if $m(x)$ interpolates $f(x)$ at the points of $\mathcal{Y}$, then, for all $x$,

$$m(x) = \sum_{j=0}^{p} f(y_i)\ell_j(x). \tag{3.2}$$

For more details and other properties of Lagrange polynomials see Section 3.2 in [9].

From (3.2) it is relatively easy to derive the relation between the value of Lagrange polynomials and the accuracy of the interpolation at a given point. Specifically,

$$|m(x) - f(x)| \leq M \sum_{j=0}^{p} \|y_j - x\|^{d+1} |\ell_j(x)|, \tag{3.3}$$

where $M$ is a constant which depends only on the Lipschitz constant of $\nabla f(x)$. See [2] for a comprehensive treatment of interpolation error bounds expressed via Lagrange polynomials.

It is clear from (3.3) that the absolute values of the Lagrange polynomials is the key indicator of the "geometry" of the interpolation set that we discussed above. If we consider now a ball $\mathcal{B}$ of radius $\Delta$ which contains $\mathcal{Y}$, then, in general, the smaller the maximum absolute value of the Lagrange polynomials on $\mathcal{B}$ the better $m(x)$ approximates $f$ on $\mathcal{B}$. In fact we can see that the bound (3.3) is similar to the Taylor expansion. However, unlike in the Taylor expansion case, to obtain a better agreement between $f(x)$ and $m(x)$ one has to consider not only the ball of a smaller radius, but a new interpolation set $\mathcal{Y}$ which fits into such a ball, while the maximum absolute value of the Lagrange polynomials has to remain bounded by the same constant.

We will make use of the following concept (borrowed from [7] and [9]) of $\Lambda$–poisedness of an interpolation set.

**Definition 3.2.** *Let $\Lambda > 0$ and a set $\mathcal{B} \in \mathbf{R}^n$ be given. A poised set $\mathcal{Y} = \{y_0, y_1, \ldots, y_p\}$ is said to be $\Lambda$–poised in $\mathcal{B}$ if and only if, for the basis of Lagrange polynomials associated with $\mathcal{Y}$, one has that*

$$\Lambda \geq \max_{j=0,\ldots,p} \max_{x\in\mathcal{B}} |\ell_j(x)|.$$

The following lemma [20] (see also [9]) is central to the use of Lagrange polynomials in geometry maintenance that we discuss here.

**Lemma 3.3.** *Given a closed bounded domain $\mathcal{B}$, any initial interpolation set $\mathcal{Y} \in \mathcal{B}$ and a constant $\Lambda > 1$, consider the following procedure: find $j \in \{0, \ldots, p\}$ and a point $x \in \mathcal{B}$ such that $|\ell_j(x)| \geq \Lambda$ (if such a point exists), and replace $y_j$ by $x$ to obtain a new set $\mathcal{Y}$. Then this procedure terminates after a finite number of iterations with a sample set which is $\Lambda$–poised in $\mathcal{B}$.*

### 3.2 Fully linear and quadratic models

We have observed that an interpolation model based on a $\Lambda$-poised sample set provides a Taylor-like approximation of the objective function. For the purposes of the algorithmic framework we may

want to abstract from the specifics of Lagrange polynomials and interpolation models (we will return to them later).

In [8] and [9] general concepts of fully-linear and fully-quadratic models were introduced.

Loosely speaking we call a model $m(x)$ to be a *fully-linear* model of $f(x)$ in $\mathcal{B}(x,\Delta)$ if

– the error between the gradient of the model and the gradient of the function satisfies

$$\|\nabla f(y) - \nabla m(y)\| \le \kappa_{eg}\,\Delta, \quad \forall y \in \mathcal{B}(x;\Delta),$$

and

– the error between the model and the function satisfies

$$|f(y) - m(y)| \le \kappa_{ef}\,\Delta^2, \quad \forall y \in \mathcal{B}(x;\Delta),$$

with constants $\kappa_{ef}$ and $\kappa_{eg}$ independent of $y$.

Analogously we call $m(x)$ *fully quadratic* in $\mathcal{B}(x,\Delta)$ if

- the error between the Hessian of the model and the Hessian of the function satisfies

$$\|\nabla^2 f(y) - \nabla^2 m(y)\| \le \kappa_{eh}\,\Delta, \quad \forall y \in \mathcal{B}(x;\Delta),$$

- the error between the gradient of the model and the gradient of the function satisfies

$$\|\nabla f(y) - \nabla m(y)\| \le \kappa_{eg}\,\Delta^2, \quad \forall y \in \mathcal{B}(x;\Delta),$$

and

- the error between the model and the function satisfies

$$|f(y) - m(y)| \le \kappa_{ef}\,\Delta^3, \quad \forall y \in \mathcal{B}(x;\Delta),$$

with constants $\kappa_{ef}$, $\kappa_{eg}$ and $\kappa_{eh}$ independent of $y$.

It is then required that there exists an algorithm which in a finite effort either certifies that a given model is fully-linear (or fully-quadratic), on a given $\mathcal{B}(x,\Delta)$ and for given constants, or constructs such a model, if it exists.

It is then shown in [9] that by means of Lagrange polynomials and Lemma 3.3 or other similar mechanisms such algorithms exist for polynomial interpolation. The proof of the fact relies on Lemma 3.3. In the framework that we describe below the abstract concept of fully-linear and fully-quadratic models is utilized.

### 3.3 A trust-region framework

Let $m(x)$ define a local model of the objective function $f(x)$ of (3.1) in the framework of trust-region algorithms. Such algorithms are iterative and build, around an iterate $x_k$, a model $m_k(x_k+s)$ of the objective function which is assumed to represent this latter function sufficiently well in a "trust region" $\mathcal{B}(x_k,\Delta_k)$, where $\Delta_k$ is known as the radius of the trust region. The model is then minimized (possibly approximately) in $\mathcal{B}(x_k,\Delta_k)$ to define a trial step $x_k^+$, and the value $f(x_k^+)$ is then computed. If this value achieves (a fraction of) the reduction from $f(x_k)$ which is anticipated on the basis of the model reduction $m_k(x_k) - m_k(x_k^+)$, then the trial point is accepted as the new iterate, the model is updated and the trust-region radius is possibly increased: this is a "successful iteration". If, on the contrary, the reduction in the objective function is too small compared to the predicted one, then the trial point is rejected and the trust-region radius is decreased: this is an unsuccessful iteration. (See [3] for an extensive coverage of trust-region algorithms.)

Thus we can roughly describe the following trust region algorithmic framework.

**Algorithm 3.1.**

**Step 0: Initialization.** *Choose an initial trust region radius $\Delta_0$, an initial poised interpolation set $\mathcal{Y}_0$ and a starting point $x_0$. This interpolation set defines an (at most quadratic) interpolation model $m_0$ around $x_0$. Chose appropriate constants for the description that follows.*

**Step 1: Criticality Step.** *If the current model gradient is much smaller than the radius of the ball containing the sample set and $x_k$ then recompute a fully-linear model based on another sample set which is closer to $x_k$ until the radius and the gradient are comparable. Set the trust region radius to be comparable to the size of the gradient as well.*

**Step 2: Compute a trial point.** *Compute $x_k^+$ such that $\|x_k^+ - x_k\| \le \Delta_k$ and $m_k(x_k^+)$ is "sufficiently small compared to $m_k(x_k)$".*

**Step 3: Evaluate the objective function at the trial point.** *Compute $f(x_k^+)$ and*

$$\rho_k = \frac{f(x_k) - f(x_k^+)}{m_k(x_k) - m_k(x_k^+)}.$$

**Step 4: Define the next iterate.**

**Step 4a: Successful iteration.** *If $\rho_k \ge \eta$, define $x_{k+1} = x_k^+$ and choose $\Delta_{k+1} \ge \Delta_k$. Obtain $\mathcal{Y}_{k+1}$ by exchanging one of the interpolation points with $\{x_k^+\}$.*

**Step 4b: Unsuccessful iteration.** *If $\rho_k < \eta$, then define $x_{k+1} = x_k$ and reduce $\Delta_k$ by a constant factor if $m_k(x)$ is fully-linear, otherwise keep $\Delta_k$ the same. Possibly update $\mathcal{Y}_{k+1}$ to include $x_{k+1}$.*

**Step 5: Update the sample set and the model.** *If the model $m_k$ is not fully-linear, then make at least one step of the model improving algorithm. Increment $k$ by one and go to Step 1.*

This algorithmic framework is theoretical and leaves many options open. For instance, what we meant by "sufficiently small' compared to $m_k(x_k)$" (in Step 1) is not specified. The first order convergence analysis merely requires that

$$m_k(x_k) - m_k(x_k^+) \ge \kappa_c \|g_k\| \min\left[\frac{\|g_k\|}{1+\|H_k\|}, \Delta_k\right],$$

where we define $g_k = \nabla m_k(x_k)$ and $H_k = \nabla^2 m_k(x_k)$, and where $\kappa_c$ is some constant in $(0,1)$ – the condition well-known in trust-region analysis under the name of "Cauchy condition".

The trust region maintenance is flexible in Step 4, while in Step 5, all is required is that an algorithm is used which can construct a fully-linear model in a finite number of steps.

The fully-quadratic models can also be used if $\nabla^2 f(x)$ is Lipschitz continuous and second order conditions are used in Steps 1 and 2 (see [3] and [8] for formal statements and details of second order conditions).

In [8] it is shown that an algorithm based on this framework (with some additional flexibility) converges to first order stationary point(s) in the case of fully-linear models and to second-order stationary point(s) in the case of fully-quadratic models and second order conditions in Steps 1 and 2.

The theory provides foundation for some existing and possible future model-based DFO algorithms, but it strongly depends on the model improvement steps (in Steps 1 and 5) where some (hopefully not many) extra sample points need to be introduced and their function values computed. Although such extra points are also computed in practical implementations of the DFO algorithms (such as NEWUOA [19] and DFO [6]) the question still remained how necessary these extra steps are.

## 4 Is it necessary to consider geometry?

### 4.1 Geometry-free framework

In particular, at the same time as the development of the convergence theory for model-based DFO methods, Fasano, Nocedal and Morales in [11] proposed an implementation of a model-based trust region method which avoided all geometry considerations entirely. Here is rough outline of the algorithm they proposed:

**Algorithm 4.1.**

**Step 0: Initialization.** *Choose an initial trust region radius $\Delta_0$, an initial poised interpolation set $\mathcal{Y}_0$ and a starting point $x_0$. This interpolation set defines an (at most quadratic) interpolation model $m_0$ around $x_0$. Chose appropriate constants.*

**Step 1: Compute a trial point.** *Compute $x_k^+$ such that $\|x_k^+ - x_k\| \leq \Delta_k$ and $m_k(x_k^+)$ is "sufficiently small compared to $m_k(x_k)$".*

**Step 2: Evaluate the objective function at the trial point.** *Compute $f(x_k^+)$ and*

$$\rho_k = \frac{f(x_k) - f(x_k^+)}{m_k(x_k) - m_k(x_k^+)}.$$

**Step 3: Define the next iterate.**

**Step 3a: Successful iteration.** *If $\rho_k \geq \eta$, define $x_{k+1} = x_k^+$ and choose $\Delta_{k+1} \geq \Delta_k$. Define the new interpolation set $\mathcal{Y}_{k+1}$ by including $x_k^+$ and by removing from $\mathcal{Y}_k$ the point $y_k$ which is the furthest away from $x_k$.*

**Step 3b: Unsuccessful iteration.** *If $\rho_k < \eta$, then define $x_{k+1} = x_k$ and reduce $\Delta_k$ by a constant factor. If $x_k^+$ is closer to $x_k$ than any of the points in $\mathcal{Y}_k$ then replace the furthest point in $\mathcal{Y}_k$ with $x_k^+$.*

**Step 4: Update the sample set and the model.** *If $\mathcal{Y}_k$ changed, update the model $m_k$. Increment $k$ by one and go to Step 1.*

This algorithm computes only one sample point per iteration and each such point is computed in the hope of reducing the objective function, hence it seems the least wasteful in terms of function evaluations. Indeed the computational results produced by the implementation were quite encouraging. However we have to note that the results were obtained by using a complete quadratic interpolation models, which require $(n+1)(n+2)/2$ sample points. If is possible that even if quadratic models become quite inaccurate they still contain valuable curvature information (see the column by J. Nocedal in this issue).

So do we need to be concerned about the geometry of sample sets or not?

### 4.2 Why considering geometry is necessary

In [20] it is shown, by two examples, that some geometry considerations are necessary in order to guarantee global convergence. Below we present one of the examples on which the algorithm proposed in [11] (and discussed in the previous subsection) produces a nonpoised set of points and converges to a nonstationary point.

Consider the following starting set of interpolation points:

$$\mathcal{Y} = \left\{ \begin{pmatrix} 11 \\ 1 \end{pmatrix}, \begin{pmatrix} 11 \\ 0 \end{pmatrix}, \begin{pmatrix} 10 \\ -1 \end{pmatrix}, \begin{pmatrix} 10 \\ 1 \end{pmatrix}, \begin{pmatrix} 10 \\ 0 \end{pmatrix}, \begin{pmatrix} 9 \\ 0 \end{pmatrix} \right\}.$$

This set is $\Lambda$–poised, in a ball of radius $2$ around $x = (10,0)^T$, with $\Lambda < 2.25$. Assume that we are given a function $f(x)$ for $x = (x_1, x_2)^T$ with the following function values on $\mathcal{Y}_0$:

$$\{121 + \alpha,\ 121,\ 100 + \alpha,\ 100 + \alpha,\ 100,\ 81\},$$

for some fixed $\alpha > 0$. Also assume that along the $x_2 = 0$ subspace the function $f(x)$ reduces to $x_1^2$ and has a minimum at $x_1 = 0$. For instance the simple function

$$f(x) = \begin{cases} x_1^2 + \alpha(x_2^2 + (10 - x_1)x_2) & \text{if } x_1 < 10; \\ x_1^2 + \alpha x_2^2 & \text{if } x_1 \geq 10, \end{cases}$$

has such properties. Note that this function has a discontinuous Hessian, however, $\nabla f(x)$ is Lipschitz continuous, so convergence to a first order stationary point is possible. Also observe that it is possible to construct a function in $C^2$ with the same properties as $f(x)$.

Now let us consider a quadratic model based on $\mathcal{Y}$. It is easy to see that the model is

$$m(x) = x_1^2 + \alpha x_2^2.$$

Choose now a trust region of radius $\Delta = 2$ centered around $y_4 = (10,0)^T$.

The iterates produced by the algorithm are shown in Figure 1. At the end the interpolation set is completely aligned with the direction $x_2 = 0$ and the model degenerates into $m(x) = x_1^2$. The algorithm then terminates at the point $x = (0,0)$, which is obtained at the next iteration and which is a non-stationary point for the original function $f(x)$.

We see here that, if the gradient of the model converges to zero, it does not imply that so does the gradient of the true function, unless the poisedness of the interpolation set is maintained.

## 5 The new algorithm

It is indeed necessary to consider geometry of the sample set to guarantee convergence of the trust-region model-based DFO methods, however, it turns out that it is not necessary to compute extra sample points unless the gradient of the model becomes small.

The final algorithm that we present here relies on a remarkable "self-correcting" property of the trust-region framework. Recall the bound (3.3). Assume that the trust region step is not successful, this implies that the error $|f(x_k^+) - m(x_k^+)|$ is relatively large (if it were not, then the good agreement between the model reduction and the objective function reduction would have caused the step to be successful). Due to (3.3) $|f(x_k^+) - m(x_k^+)|$ can be relatively large only if either one of the $||x_k^+ - y_i||$ is relatively large or if one of the values $|\ell_i(x_k^+)|$ is relatively large, which in turn means that replacing one of the $y_i$'s by $x_k^+$ will improve the sample set (see Lemma 3.3). In [20] this intuition is supported by rigorous derivation and the following algorithm is proposed.

**Algorithm 5.1.**

**Step 0: Initialization.** *Choose an initial trust region radius $\Delta_0$, an initial poised interpolation set $\mathcal{Y}_0$ and a starting point $x_0$. This interpolation set defines an (at most quadratic) interpolation model $m_0$ around $x_0$. Chose appropriate constants.*

**Step 1: Criticality Step.** *If the current model gradient is smaller than some threshold $\epsilon_k$ then build a fully-linear model based on a sample set which is sufficiently close to $x_k$ (so that the interpolation radius and the gradient are comparable). Set the trust region radius to be comparable to the size of the gradient as well. Decrease $\epsilon_k$ by a constant factor.*

**Step 2: Compute a trial point.** *Compute $x_k^+$ such that $\|x_k^+ - x_k\| \leq \Delta_k$ and $m_k(x_k^+)$ is "sufficiently small compared to $m_k(x_k)$".*

**Step 3: Evaluate the objective function at the trial point.** *Compute $f(x_k^+)$ and*

$$\rho_k = \frac{f(x_k) - f(x_k^+)}{m_k(x_k) - m_k(x_k^+)}.$$

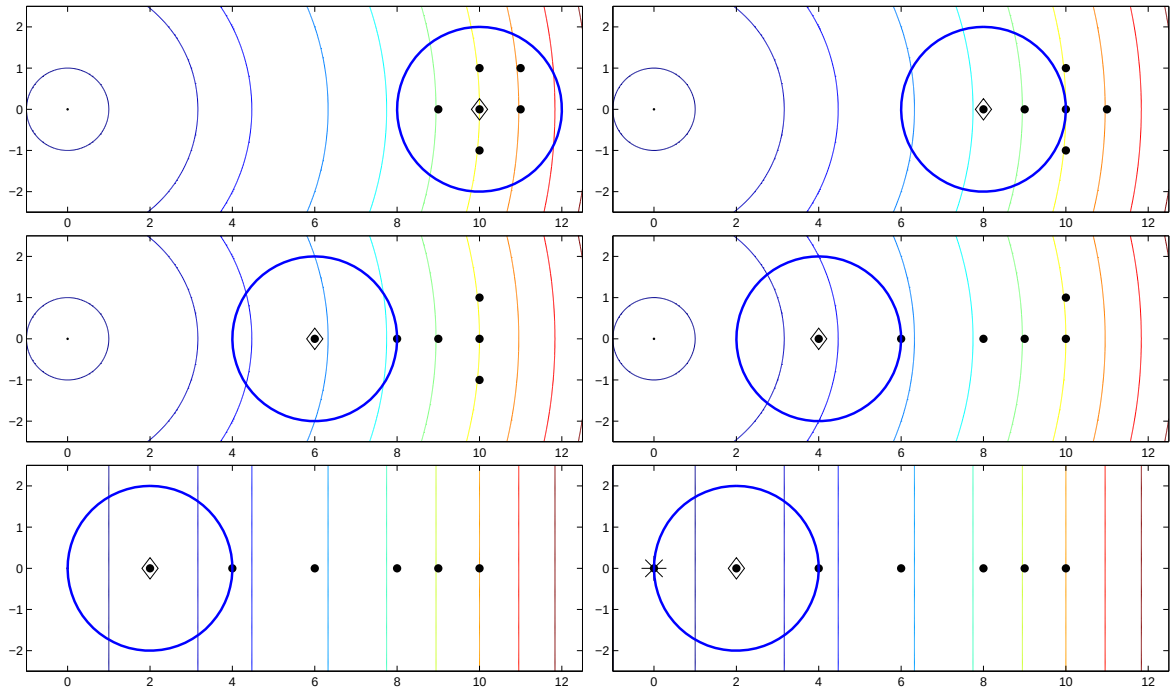


Figure 1. *From left to right and top to bottom: the successive iterates of the algorithm on the associated models, where the current iterate is marked by a diamond and surrounded by its circular-shaped trust region. The final convergence point is indicated by a star.*

**Step 4: Define the next iterate.**

**Step 4a: Successful iteration.** *If* $\rho_k \geq \eta$, *define* $x_{k+1} = x_k^+$ *and choose* $\Delta_{k+1} \geq \Delta_k$. *Update the interpolation set to obtain* $\mathcal{Y}_{k+1}$ *by swapping* $x_k^+$ *with the point* $y_{j,k}$ *in* $\mathcal{Y}_k$ *which is either far away or for which* $|\ell_{k,j}(x_k^+)|$ *is the largest.*

**Step 4b: Unsuccessful iteration.** *If* $\rho_k < \eta$, *then define* $x_{k+1} = x_k$ *and*

*(i) if there is a far away point in* $\mathcal{Y}$ *then replace it with* $x_k^+$,

*(ii) if for some* $j$ $|\ell_{k,j}(x_k^+)|$ *is larger than some fixed* $\Lambda > 1$, *replace* $y_{j,k}$ *with* $x_k^+$,

*(iii) otherwise reduce* $\Delta_k$ *by a constant factor.*

**Step 5: Update the sample set and the model.** *If the interpolation set* $\mathcal{Y}_k$ *changed, then update the model* $m_k$. *Increment* $k$ *by one and go to Step 1.*

In [20] the detailed mechanism of Step 4 is derived in a way which guarantees that there can only be a finite number of consecutive unsuccessful steps while the model gradient is bounded away from zero. The proof relies on Lemma 3.3. Using this fact it is then shown (under the usual and reasonable conditions) that this algorithm has at least one limit point which is the first order stationary point for $f(x)$.

It is important to note that without Step 1 this algorithm will also fail on the example in the previous section, since the outcome of the example does not depend on the order in which interpolation points are removed from the interpolation set.

The maintenance of Lagrange polynomials does not add extra computational cost since it is the same as the cost of maintaining the quadratic model [16]. This means that aside from the iterations which invoke Step 1, each iteration of this algorithms is essentially the same in terms of computational cost as the steps of algorithm in [11]. We also note that in fact this new algorithm is the closest theoretically convergent algorithm to the practical implementations in [19] and [6] that exist so far.

It remains to be seen if stronger theoretical results can be obtained for this (or a similar) algorithm.

Katya Scheinberg, Department of Industrial Engineering and Operations Research, Columbia University, New York. katyascheinberg@gmail.com

## References

[1] M. A. Abramson and C. Audet. Convergence of mesh adaptive direct search to second-order stationary points. *SIAM Journal on Optimization*, 17:606–619, 2006.

[2] P. G. Ciarlet and P. A. Raviart. General Lagrange and Hermite interpolation in $R^n$ with applications to finite element methods. *Arch. Ration. Mech. Anal.*, 46:177–199, 1972.

[3] A. R. Conn, N. I. M. Gould, and Ph. L. Toint. *Trust-Region Methods*. Number 01 in MPS-SIAM Series on Optimization. SIAM, Philadelphia, USA, 2000.

[4] A. R. Conn, K. Scheinberg, and Ph. L. Toint. On the convergence of derivative-free methods for unconstrained optimization. In A. Iserles and M. Buhmann, editors, *Approximation Theory and Optimization: Tributes to M. J. D. Powell*, pages 83–108, Cambridge, England, 1997. Cambridge University Press.

[5] A. R. Conn, K. Scheinberg, and Ph. L. Toint. Recent progress in unconstrained nonlinear optimization without derivatives. *Mathematical Programming, Series B*, 79(3):397–414, 1997.

[6] A. R. Conn, K. Scheinberg, and Ph. L. Toint. A derivative free optimization algorithm in practice. Proceedings of the 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, St. Louis, Missouri, September 2-4, 1998.

[7] A. R. Conn, K. Scheinberg, and L. N. Vicente. Geometry of interpolation sets in derivative free optimization. *Mathematical Programming*, 111:141–172, 2008.

[8] A. R. Conn, K. Scheinberg, and L. N. Vicente. Global convergence of general derivative-free trust-region algorithms to first and second order critical points. *SIAM Journal on Optimization*, (to appear), 2009.

[9] A. R. Conn, K. Scheinberg, and L. N. Vicente. *Introduction to Derivative-free Optimization*. MPS-SIAM Optimization series. SIAM, Philadelphia, USA, 2009.

[10] A. R. Conn and Ph. L. Toint. An algorithm using quadratic interpolation for unconstrained derivative free optimization. In G. Di Pillo and F. Gianessi, editors, *Nonlinear Optimization and Applications*, pages 27–47, New York, 1996. Plenum Publishing.

[11] G. Fasano, J. Nocedal, and J.-L. Morales. On the geometry phase in model-based algorithms for derivative-free optimization. *Optimization Methods and Software*, (to appear), 2009.

[12] J. A. Nelder and R. Mead. A simplex method for function minimization. *Comput. J.*, 7:308–313, 1965.

[13] M. J. D. Powell. A direct search optimization method that models the objective and constraint functions by linear interpolation. In S. Gomez and J. P. Hennart, editors, *Advances in Optimization and Numerical Analysis, Proceedings of the Sixth Workshop on Optimization and Numerical Analysis, Oaxaca, Mexico*, volume 275, pages 51–67, Dordrecht, The Netherlands, 1994. Kluwer Academic Publishers.

[14] M. J. D. Powell. A direct search optimization method that models the objective by quadratic interpolation. Presentation at the 5th Stockholm Optimization Days, Stockholm, 1994.

[15] M. J. D. Powell. A quadratic model trust region method for unconstrained minimization without derivatives. Presentation at the International Conference on Nonlinear Programming and Variational Inequalities, Hong Kong, 1998.

[16] M. J. D. Powell. On the Lagrange functions of quadratic models that are defined by interpolation. *Optim. Methods Softw.*, 16:289–309, 2001.

[17] M. J. D. Powell. On the use of quadratic models in unconstrained minimization without derivatives. Technical Report NA2003/03, Department of Applied Mathematics and Theoretical Physics, Cambridge University, Cambridge, England, 2003.

[18] M. J .D. Powell. Least Frobenius norm updating of quadratic models that satisfy interpolation conditions. *Mathematical Programming, Series B*, 100(1):183–215, 2004.

[19] M. J. D. Powell. Developments of NEWUOA for minimization without derivatives. *IMA Journal of Numerical Analysis*, 28(4):649–664, 2008.

[20] K. Scheinberg and Ph. L. Toint. Self-correcting geometry in model-based algorithms for derivative-free unconstrained optimization. Technical report, submitted, 2009.

[21] V. Torczon. On the convergence of the multidirectional search algorithm. *SIAM Journal on Optimization*, 1:123–145, 1991.

[22] V. Torczon. On the convergence of pattern search algorithms. *SIAM Journal on Optimization*, 7:1–25, 1997.

[23] P. Tseng. Fortified-descent simplicial search method: a general approach. *SIAM Journal on Optimization*, 10:269–288, 1999.

# Discussion column

Jorge Nocedal

## Finding the middle ground between first and second-order methods

In the last few years, we have witnessed the emergence of first-order methods for a variety of nonlinear optimization applications. The advocacy of first-order methods is, however, in stark contrast with much of the algorithmic practice of the last 30 years that has emphasized methods based on quadratic models to achieve faster convergence. Therefore, it is reasonable to ask whether this shift in emphasis is well justified.

Several arguments have been advanced in favor of first-order methods.

1. For very large and data-intensive problems, inexpensive first-order methods are more efficient than more rapidly convergent, but more expensive, methods. Furthermore, in cases where approximate solutions are adequate the benefits of second-order methods are not realized since the optimization is terminated early.
2. In some applications, derivatives are not available, and approximating Hessian matrices is very costly in methods for derivative-free optimization.
3. For non-smooth problems, quadratic approximations may not be appropriate.
4. One can establish complexity results for first-order methods on certain challenging problem classes. To establish similar results for higher-order methods requires unrealistic assumptions.
5. First-order methods are more than adequate for problems that contain uncertainty in the data.

Undoubtedly, there are some cases when first-order methods are the right tool for the job. But in many contexts, some of the arguments given above are not well justified and lead to algorithms that are unnecessarily slow. The key observation of this note is that *simple* quadratic models that do not attempt to accurately approximate the Newton model often give rise to very attractive algorithms for many of the situations listed above. Indeed, the judicious use of selective second-order information can bring dramatic savings in computing time.

The article by Katya Scheinberg in this issue deals with derivative-free optimization, one of the areas in which the use of (low quality) quadratic models has proved surprisingly effective. Her article explores the limits of inaccuracy in model-based methods for derivative-free optimization, a class of methods pioneered by Powell and by Scheinberg and her collaborators; see [1]. Numerical experience has shown that quadratic models give rise to much more effective methods than linear models, even though they require $O(n^2)$ vs $O(n)$ function values to define the model via interpolation. Furthermore, Powell [10] has recently proposed a framework for *updating* quadratic models using only $O(n)$ interpolation points, which allows the method to solve much larger problems than in the past. Needless to say, such an approach does not aim to generate a good approximation of the Hessian matrix – and there is no hope of obtaining superlinear convergence – but the method constitutes a superior approach for derivative-free optimization. Some remarkable numerical experiments by Moré and Wild [8] indicate that Powell's model-based method is very effective (and clearly outperforms a leading pattern search method) even for certain classes of *non-smooth* problems. This efficiency is achieved in spite of the very low accuracy of their quadratic approximations. As Scheinberg discusses in this issue of Optima, one needs to impose only minimal quality controls to promote convergence and ensure good performance.

Equally surprising is the recent study by Lewis and Overton [5] on general-purpose methods for the minimization of locally Lipschitz nonsmooth functions. They observe that the BFGS quasi-Newton method is far more effective than more conservative techniques, such as bundle methods. Since locally Lipschitz functions are differentiable almost everywhere, the BFGS iteration (with an appropriate line search) is normally well defined and is able to approximate solutions even when they occur at a point of nondifferentiability [1]. As is the case in Powell's method for derivative-free optimization, the quadratic models become extremely ill conditioned, but this does not prevent the methods from moving along fruitful search directions. Lewis and Overton do not provide convergence results (except for very simple special cases) but offer good insights; for example they report that the BFGS matrix often provides a good approximation of the so-called U and V spaces associated with the objective function. One cannot yet claim that the BFGS method represents a general-purpose algorithm for non-smooth optimization because it typically breaks down close to the solution and is therefore unable to provide a certificate of optimality. Nevertheless, the renewed interest in the use of second-order information is yet to be fully developed in non-smooth optimization.

Two of the most popular methods for smooth large-scale optimization, the inexact (or truncated) Newton method and limited memory BFGS method, are typically implemented so that the rate

of convergence is only linear. They are good examples of algorithms that fall between first and second-order methods. Let me mention three specific application areas where significant progress has been made by designing new methods of this kind.

Rigid body simulations for computer game simulations often lead to linear complementarity problems (LCP) that must be solved very quickly because graphics operate at 60 frames per second. Unlike studio animation movies, computer game animations need not be of very high quality, and therefore the solution of the linear complementarity problem is terminated quite early. Game developers typically use the projected Gauss-Seidel (or projected SOR) method to compute very approximate solutions of the linear complementarity problems. This prototypical first-order method has been advocated by the gaming community because the use of second-order methods is not practical (interestingly, interior-point methods are not well suited in this case).

Kocvara et al. [3] and Morales et al. [7] have shown that, by interlacing a subspace improvement iteration with the projected Gauss-Seidel iteration, it is possible to compute very accurate solutions in less time. When the LCP is symmetric, the subspace improvement space amounts to the minimization of an associated quadratic program over a set of active free variables, and [7] shows that it is effective to use exact second order information on this subspace. The key observation is that the subspace improvement phase greatly accelerates the identification of the optimal constraints – it does not simply provide a higher rate of convergence once this identification has been made. These advances might not be altogether surprising given that Polak [9] demonstrated long ago that the gradient projection method benefits greatly from a subspace minimization phase. Nevertheless, it is often not straightforward to translate general design principles from one context to another one.

In fact, essentially the same approach that proved effective in computer game simulations has recently been applied by Wen et al. [11] in compressive sensing applications, and by Feng et al. [2] in the pricing of American options. In these two papers the subspace minimization phase uses an iterative approach (CG or BFGS for compressive sensing, GMRES with ILU preconditioner, for options pricing). Significant speedups are obtained with respect to first-order methods.

It is difficult to explain precisely why a minimal use of second-order information can bring substantial benefits. In the case of BFGS for nonsmooth optimization, approximate Hessians provide much needed information of curvature of functions, but also, more importantly, information about changes in the function due to discontinuities in first-order derivatives. All the new methods I have mentioned in this column, can be seen as occupying a middle ground between first and second-order methods. This is fertile territory.

Jorge Nocedal, Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, USA. nocedal@eecs.northwestern edu

## References

[1] A. R. Conn, K. Scheinberg, and L. Vicente. *Introduction to derivative-free optimization*. SIAM, Philadelphia, USA, 2009.

[2] L. Feng, J.L. Morales, J. Nocedal, and V. Linetsky. On the solution of complementarity problems arising in American options pricing. Technical Report OTC 09/02, Northwestern University, 2009.

[3] M. Kocvara and J. Zowe. An iterative two-step algorithm for linear complementarity problems. *Numerische Mathematik*, 68:95–106, 1994.

[4] C. Lemaréchal. Numerical experiments in nonsmooth optimization. In E. A. Nurminski, editor, *Progress in nondifferentiable optimization*, volume 26, pages 61–84, Laxenburg, Austria, 1982. The Interface Foundation of North America.

[5] A.S. Lewis and M.L. Overton. Nonsmooth optimization via BFGS. Technical report, New York University, 2009. submitted to SIAM Journal on Optimization.

[6] L. Luksan and J. Vlcek. Globally convergent variable metric methods for convex nonsmooth unconstrained optimization. *Journal of Optimization Theory and Applications*, (102):593–613, 1999.

[7] J.L. Morales, J. Nocedal, and M. Smelyanskiy. An algorithm for the fast solution of linear complementarity problems. *Numerische Mathematik*, 111(2):251–266, 2008.

[8] J. J. Moré and S. M. Wild. Benchmarking derivative-free optimization algorithms. *SIAM J. Optimization*, 20(1):172–191, 2009.

[9] B. T. Polyak. The conjugate gradient method in extremal problems. *U.S.S.R. Computational Mathematics and Mathematical Physics*, 9:94–112, 1969.

[10] M. J. D. Powell. Least Frobenius norm updating of quadratic models that satisfy interpolation conditions. *Mathematical Programming*, 100(1):183–215, 2004.

[11] Z. Wen, W. Yin, D. Goldfarb, and Y. Zhang. A fast algorithm for sparse reconstruction based on shrinkage, subspace optimization and continuation. Technical Report CAAM Technical Report TR09-01, Rice University, 2009.

Steve Wright

# MPS Chair's Column

April 16, 2009. I welcome the new Optima! As would be evident by the time you read these words, our society's newsletter is now being designed and produced under new arrangements. The designer and typesetter is Christoph Eyrich, who is also in charge of the design and production of the German Mathematical Society (DMV) newsletter. These new arrangements represent the next major step in our drive to renew Optima by publishing it on a regular schedule, speeding up the production/distribution process, and re-conceiving the content. I thank the many people in MPS who have contributed to this process, most especially editor Andrea Lodi.

It's time to recognize the tremendous contributions of Don Hearn, who was the founding editor of Optima (in 1980) and who has served continuously as editor and then publisher since that time. Don and his publication team at U. Florida have played a central role in sustaining Optima (and thus MPS) over three decades, and we thank them for their dedicated service.

All past issues of Optima can be found on the MPS web site http://www.mathprog.org. A column by Michael Held (then Publications Committee Chair) in Issue 1 (1980) explains the origins of Optima, which grew out of discussions among the MPS leadership of the time, including MPS chair Phil Wolfe, George Nemhauser and Michael Powell.

The new MPS web site went live in January 2009 after an extended and careful redesign. The new site is much easier to maintain and (we hope) easier to navigate. Special thanks to webmaster Marc Pfetsch for his initiative and hard work during this process. If you have optimization-themed photos for including in the album, or any other comments, on the site, please contact Marc.

Also newly available on http://www.mathprog.org are the issues of the COAL Newsletter published between 1979 and 1993. Thanks to Trond Steihaug for supplying the scans of these newsletters, which were influential during a key period in algorithm and software development for optimization. COAL – the Committee on Algorithms – was formerly a standing committee of MPS.

Mathematical Programming Studies, the predecessor of Mathematical Programming, Series B, which was published in 31 volumes between 1974 and 1987, is now available free to MPS members on the SpringerLink web site. If you log in with your personal MPS identifier for SpringerLink, you should have full-text access to all papers. These volumes contain many influential papers in the development

of our field, along with interesting historical information about leading figures in mathematical programming, in the dedications of some volumes. To get to the right page on SpringerLink, you can follow the link from the mathprog.org web site, or else use the "Find" box in SpringerLink. Speaking of SpringerLink, it is a good idea to add the pages for Mathematical Programming, Mathematical Programming Computation, and Mathematical Programming Studies to your list of saved items on this site, for easy access each time you log in.

We look forward to the 20th ISMP (August 23–28, 2009; www.ismp2009.org), which returns to Chicago for the first time since the "Zeroth ISMP" in 1949 (sixty years ago!) and the 4th ISMP in 1962. This symposium will be held at a particularly exciting time for our field. Optimization is playing a vital role in more and more areas of science and engineering, and awareness continues to grow of the key contributions that optimization can make to many interdisciplinary projects. Our field is revitalized by the new paradigms and formulations that arise in these applications, which are often extremely challenging because of their size, their complexity, the need for approximate solutions in real time, and the need to incorporate risk and uncertainty in the models. Recent failures in economic/financial systems present us with new opportunities to influence public policy through better risk models and better algorithms. We face the challenges of finding credible collaborators in the financial and economic fields, and of interacting with decision-makers to identify policies that are politically and socially feasible, as well as near-optimal by some measure.

As I write, abstract submission for ISMP 2009 has just closed and indications are that attendance will be high, despite challenging economic times. If you submitted an abstract, please remember to register by the author registration deadline of May 29 to ensure that your talk is scheduled. I also recommend to book accommodation early (see the conference web site for information) and to book early for the banquet if you wish to attend. The banquet will be held on Wednesday evening at the Field Museum, Chicago's famous natural history museum, and tickets are strictly limited.

Committees have been working hard to select winners of the prizes to be awarded during the ISMP opening ceremony, to be held on August 23 at Chicago's Orchestra Hall. Please also reserve the Tuesday evening of ISMP, after sessions conclude, for the MPS business meeting. Here, we will have a membership vote on the new MPS constitution, present the new officers and council, and announce the location of ISMP 2012.

Steve Wright, MPS Chair

## New Constitution and Bylaws for MPS

A few years ago, MPS was advised that its bylaws were not quite in the standard form expected for non-profit organizations under US tax law. The bylaws had been drafted with the help of an attorney when the society was founded in 1970, but amendments in later years had taken place generally without legal advice. We decided to do a thorough redrafting of the bylaws, bringing them into line with current MPS practice, adding a new section to account for ICCOPT, incorporating the Prize rules, and adding precision in many places. Our aim was not merely to satisfy the legal requirements but also to provide a reference document for future MPS officers, editors, and conference organizers. Naturally, we consulted with the society's attorney to ensure that the final document could pass legal muster. We hope that it will serve the society, with minor amendments and additions as needed, for at least the next 20 years. The new bylaws were approved by vote of MPS Council on 23 Feb. 2009, and can be found on the MPS web site at www.mathprog.org.

We took this opportunity to amend the Society's constitution as well. The changes are intended to modernize and clarify the document. They are minor, but too numerous to be detailed here; I urge you to read the proposed new version which is printed below and which can also be found on the web site. Amendments to the constitution require approval by the full membership of the Society. Council recommends a vote in favour of the new constitution. A vote of the membership will be held at the MPS business meeting during ISMP, where a simple majority will suffice to approve, providing a quorum is present.

Thanks to all those who contributed to the final versions, especially David Gay, who worked on various drafts and handled the communications with our attorney.

## Constitution of the Mathematical Programming Society*

### I Name

The society is an international organization to be called "Mathematical Programming Society, Inc." It will henceforth be referred to as the Society.

### II Objectives

The objectives of the Society are the communication of knowledge of the theory, applications, and computational aspects of mathematical programming and related areas and the stimulation of their development. To realize these objectives, the Society publishes several journals, holds International Symposia and sponsors such other activities consistent with the objectives as may be directed by the Council.

### III Membership

The membership of the Society consists of individual members and of corporate members. Members join the Society by application in a form prescribed by the Council.

### IV Council

1. The elected members of the Council of the Society are the Chair, the Vice-Chair, the Treasurer, and four at-large members. The Chair of the Executive Committee, the chair of the Publications Committee, and the Editors-in-Chief of the journals shall be invited to all Council meetings and shall be included on all Council correspondence. All must be members of the Society.

2. The Chair chairs the meetings of the Council. The Council votes by majority of the elected members present, with the Chair having a casting vote.

3. The Chair will submit a report on the activities of the Society when he** relinquishes the office. This report will be published in a journal or newsletter of the Society. The Chair will chair a business meeting on the occasion of any International Symposium held during his term of office.

4. The Vice-Chair replaces the Chair whenever the necessity arises.

5. The Treasurer is responsible for the administration of the funds of the Society, as directed by the Council. The Treasurer shall make a financial report to the Society at the International Symposium held within his term of office.

6. The Editors-in-Chief of the journals are appointed by the Council subject to the terms of the contract in force with publishers

of the journals. They are responsible for implementing the directives of the Council, in the organization of the journals, and for carrying out its policy.

7. At each International Symposium there will be a combined meeting of the outgoing Council and the incoming Council. Additional meetings must be held when requested by at least three members of the Council. The place of such meetings is decided by the Chair. The Chair makes arrangements for the taking of minutes at meetings of the Council and business meetings of the Society.

8. The policies of the Council are carried out by the Executive Committee. The chair of the Executive Committee is appointed by the Council, following a nomination by the Chair, which the Council may approve or disapprove, and thereafter serves until the Chair nominates a replacement candidate for the office. The chair of the Executive Committee is responsible for executing the executive directives of the Council and for advising the Council. The Chair, Vice-Chair and Treasurer are ex-officio members of the Executive Committee. The Chair may appoint additional members of the Executive Committee, as necessary to allow the Executive Committee to carry out its purpose. Such members serve at the pleasure of the Chair.

9. The Council appoints such other committees as it finds necessary to carry out the business of the Society or to further its objectives. The Chair and the chair of the Executive Committee are ex-officio members of all such committees, except for those committees formed for purposes of determining winners of the Society's prizes.

## V International Symposia

1. International Symposia are sponsored by the Society at intervals of approximately three years. The Chair nominates and the Council elects the chair of the Organizing Committee and the chair of the Program Committee of the next International Symposium.
2. Fees for the International Symposium are fixed by the Organizing Committee, in consultation with the Chair. The Council shall adopt guidelines regarding the financial obligations between the Society and the Organizing Committee.

## VI Elections

1. In this section, the word "term" is defined to be the period from the end of one International Symposium to the end of the following International Symposium.
2. Elections for the Offices of Chair, Treasurer and the four at-large members of Council are concluded at least two months prior to each International Symposium. The elected Chair serves on Council for the two terms following his election. He is the Chair from one year after the beginning of the first term until one year after the beginning of the second term. He takes the office of Vice-Chair during the remainder of his period of service. The Treasurer takes office one year after the beginning of the term following his election and he serves until one year after the beginning of the next term. At-large members of Council serve for the term following their election. If the office of Chair becomes vacant, it is filled automatically by the Vice-Chair. The Chair, after consultation with Council, may appoint a member of the Society to fill any other office that becomes vacant until the next election. No one may serve for more than two consecutive terms as an elected at-large member of Council.
3. The Chair invites nominations for all elections, giving at least two months notice through a journal or newsletter of the Society of the closing date for the receipt of nominations. Candidates must be individual members of the Society. They may be proposed either by Council or by any six individual members of the Society. No nomination that is in accordance with the constitution may be refused, provided that the candidate agrees to stand. The Chair decides the form of the ballot. →


The 20th International Symposium on Mathematical Programming will take place August 23–29, 2009 in Chicago, Illinois. The meeting will be held at the University of Chicago's Gleacher Center and the Marriott Downtown Chicago Magnificent Mile Hotel. Festivities planned for the conference include the opening session in Chicago's Orchestra Hall, home of the Chicago Symphony Orchestra, the conference banquet at the Field Museum, Chicago's landmark natural history museum, and a celebration of the 60th anniversary of the Zeroth ISMP Symposium.

The plenary and semi-plenary speakers are
- Eddie Anderson, University of Sydney
- Mihai Anitescu, Argonne National Lab
- Stephen Boyd, Stanford University
- Friedrich Eisenbrand, EPFL
- Matteo Fischetti, University of Padova
- Lars Peter Hansen, University of Chicago
- Jong-Shi Pang, University of Illinois at Urbana-Champaign
- Pablo Parrilo, MIT
- Andrzej Ruszczynski, Rutgers
- Martin Skutella, Technische Universität Berlin
- David Shmoys, Cornell
- Eva Tardos, Cornell
- Paul Tseng, University of Washington
- Shuzhong Zhang, The Chinese University of Hong Kong

Please plan on attending the opening session on Sunday evening, where MPS prizes will be presented. Please also plan to attend the MPS business meeting on Tuesday evening, which will include an annoucement of the site of the next ISMP and a vote on the proposed new constitution.

Please keep checking the symposium web site www.ismp2009.org during the coming months, where all developments will be posted. In particular, you can register for the conference through the web site and find out about accommodation options. We urge you to book hotels as soon as possible.

One new feature at this symposium will be the daily newsletter Optima@ISMP, which will contain news about each day's events, interviews with MPS and ISMP personalities, and local information about Chicago.

### VII Secretariat

1. The Council is assisted by a Secretariat, which is supervised by the chair of the Executive Committee and Treasurer.
2. The Secretariat will keep an up-to-date list of members of the Society and a list of past and present members of the Council, with an indication of their functions.

### VIII Fees

Membership fees are fixed by Council. A member who has not paid his dues before the end of the current year will be deemed to have left the Society.

### IX Journals

Journals of the Society are distributed to all members of the Society, free of any charge additional to the membership fee, to their last known address.

### X Agents

Council may approve the payments of membership fees, or of subscription fees for the journal, in national currency, to local agents in countries where the Council, in its sole discretion, determines it is difficult for individual members to obtain convertible currency.

### XI Other activities

In addition to International Symposia, the Society may sponsor other conferences and seminars. The organization of such sponsored meetings is subject to directives by the Chair.

### XII Amendment of the Constitution

If proposed by at least ten individual members of the Society, or by vote of the Council, the constitution may be amended by a majority of individual voters, either at a business meeting of the society on the occasion of an International Symposium at which a quorum is present, or by a written ballot. Proposals must reach the Chair at least two months before the voting takes place.

### XIII Bylaws

1. To carry out the obligation as set forth in this constitution and to conduct the business of the Society, the Council shall adopt bylaws. The bylaws may be adopted, annulled, or amended by an affirmative vote of at least four members of the Council. The bylaws also may be amended by the members of the Society at any business meeting of the Society by a majority vote of those present in person or by proxy, where such meeting was called in whole or in part for that purpose and notice of such proposal was given at least thirty (30) days prior to the date of the meeting. The Council shall have the authority in its sole discretion to interpret the bylaws.
2. Council shall adopt bylaws governing elections designed to promote and maintain international representation of the Council and Executive Committee.

### Notes

* MATHEMATICAL PROGRAMMING SOCIETY is a registered trademark of the Mathematical Programming Society, Inc.

** Throughout this document, in accordance with standard English, no assumption about gender is implied by the use of a male pronoun.

## Announcement: MPS Election Results

Triennial elections have recently concluded for the Mathematical Programming Society. The elected candidates are as follows:

– Chair: Philippe Toint (University of Namur)
– Treasurer: Juan Meza (Lawrence Berkeley National Laboratory)
– Council-Members-At-Large: Jeff Linderoth (University of Wisconsin-Madison), Claudia Sagastizábal (Associação Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro), Martin Skutella (Technische Universität Berlin), and Luís Nunes Vicente (Universidade de Coimbra).

The newly elected At-Large Council members will be installed at the 20th ISMP this summer, in Chicago, and will hold office August 2009 – August 2012. The new Chair and Treasurer take office in August 2010 and will serve for the following three years. The current Chair, Stephen Wright, will be Vice-Chair during the period August 2010 – August 2012. As is readily apparent, leadership of the Society will continue to be in very good hands.

Alberto Caprara, Andrea Lodi and Katya Scheinberg

## What's new in Optima?

Optima was born in 1980 thanks to the idea of Don Hearn who has continued to ensure its existence with endless dedication and energy. At the beginning of its 30th year of life the time has come for Optima to move out of its childhood home at the University of Florida. It is moving to its new design and production site in Europe. Needless to say, we are indebted to Don for his years of service and specifically for his guidance during our own first two years as the Optima team of service – a very short period with respect to his 29 years! Nevertheless, we felt bold enough to take on the decisions necessary for the new production process of the MPS newsletter. We are glad to present Optima 79 as the first issue of the new Optima. With respect to the content we do not present big changes but we hope the new design, a splash of red and the more accurate LaTeX-based mathematical layout will be appreciated by our readers. We like to thank here the new designer Christoph Eyrich, and the Optima Committee members Steve Wright, Jon Lee, Harvey Greenberg and Mike Trick for their help in sorting out the future of the newsletter.

**IMPRINT**

***Editor:*** Andrea Lodi, DEIS University of Bologna, Viale Risorgimento 2, I-40136 Bologna, Italy. andrea.lodi@unibo.it
***Co-Editors:*** Alberto Caprara, DEIS University of Bologna, Viale Risorgimento 2, I-40136 Bologna, Italy. acaprara@deis.unibo.it
Katya Scheinberg, Department of Industrial Engineering and Operations Research, Columbia University, 500 W 120th Street, New York, NY, 10027. katyascheinberg@gmail.com
***Founding Editor:*** Donald W. Hearn

Published by the Mathematical Programming Society.
Design and typesetting by Christoph Eyrich, Mehringdamm 57 / Hof 3, 10961 Berlin, Germany. optima@0x45.de
Printed by Oktoberdruck AG, Berlin.

# MOPTA 2009

Bethlehem, PA, 19-21 August 2009

LEHIGH UNIVERSITY

## The conference

**MOPTA** is for people from both Discrete and Continuous Optimization, working on both theoretical and applied aspects.

Format: Six invited talks and many more contributed talks, spread over three days.

## Invited speakers

Ravindra K. Ahuja (UFL Gainesville)
Natalia Alexandrov (NASA)
Paul I. Barton (MIT)
John M. Mulvey (Princeton)
Pablo A. Parrilo (MIT)
Robert Weismantel (U. Magdeburg)

## Modeling competition

Organized in collaboration with AIMMS, this competition challenges groups of students from around the world to model and solve a difficult, real-world Optimization problem.

The finalists will present their work at MOPTA, where the prize for the best work will be awarded.

## The University

Located in the south of Bethlehem, PA, it is just an hour from New York City and an hour from Philadelphia.

Bus connections to Bethlehem are available from Newark, JFK, La Guardia, and Philadelphia airports.

## Organizing committee

- Tamás Terlaky (Chair)
- Pietro Belotti
- Jitamitra Desai
- Imre Pólik
- Ted Ralphs
- Larry Snyder
- Robert H. Storer
- Aurélie Thiele

## Practical information

**June 6**: abstract submission deadline
**July 1**: early registration deadline

Contact: ISE Dept, Lehigh University
200 W Packer Ave.
Bethlehem PA 18015
Phone: 610 758 3865
Email: **mopta2009@lehigh.edu**

**http://mopta.ie.lehigh.edu**

---

**Application for Membership**

I wish to enroll as a member of the Society. My subscription is for my personal use and not for the benefit of any library or institution.

☐ I will pay my membership dues on receipt of your invoice.
☐ I wish to pay by credit card (Master/Euro or Visa).

*Credit card no.* _____ *Expiration date* _____

*Family name* _____

*Mailing address* _____

_____

*Telephone no.* _____ *Telefax no.* _____

*E-mail* _____

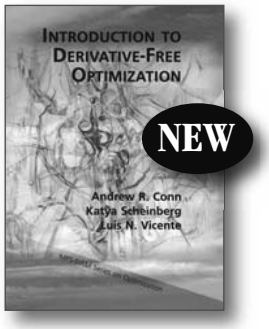
*Signature* _____

*Mail to:*
Mathematical Programming Society
3600 University City Sciences Center
Philadelphia, PA 19104-2688
USA

Cheques or money orders should be made payable to The Mathematical Programming Society, Inc. Dues for 2009, including subscription to the journal *Mathematical Programming*, are US $85. Retired are $40. Student applications: Dues are $20. Have a faculty member verify your student status and send application with dues to above address.

*Faculty verifying status* _____

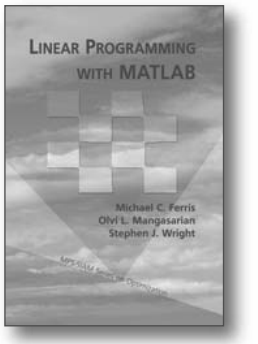
*Institution* _____




# MPS-SIAM Series on OPTIMIZATION

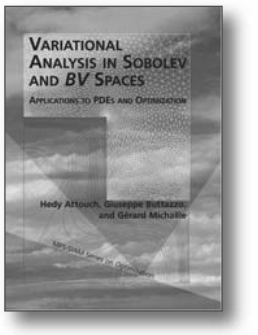
**Philippe Toint, Editor-in-Chief**
*University of Namur, Belgium*

## BOOKS IN THE SERIES INCLUDE:



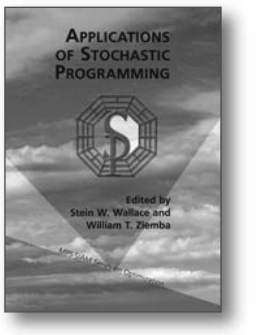

### Introduction to Derivative-Free Optimization
Andrew R. Conn, Katya Scheinberg, and Luis N. Vicente
2009 · xii + 277 pages · Softcover · ISBN 978-0-898716-68-9
List Price $73.00 · SIAM Member Price $51.10 · **Code MP08**

### Linear Programming with MATLAB
Michael C. Ferris, Olvi L. Mangasarian, and Stephen J. Wright
2007 · xii + 266 pages · Softcover · ISBN 978-0-898716-43-6
List Price $45.00 · SIAM Member Price $31.50 · **Code MP07**




### Variational Analysis in Sobolev and *BV* Spaces: *Applications to PDEs and Optimization*
Hedy Attouch, Giuseppe Buttazzo, and Gérard Michaille
2005 · xii + 634 pages · Softcover · ISBN 978-0-898716-00-9
List Price $140.00 · MPS/SIAM Member Price $98.00 · **Code MP06**

### Applications of Stochastic Programming
Edited by Stein W. Wallace and William T. Ziemba
2005 · xvi + 709 pages · Softcover · ISBN 978-0-898715-55-2
List Price $142.00 · MPS/SIAM Member Price $99.40 · **Code MP05**

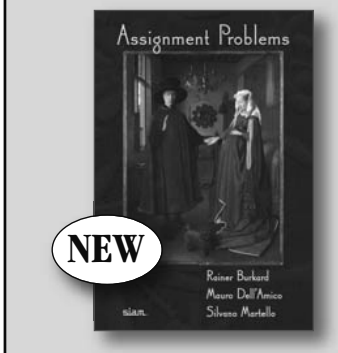



### The Sharpest Cut: *The Impact of Manfred Padberg and His Work*
Edited by Martin Grötschel
2004 · xi + 380 pages · Hardcover · ISBN 978-0-898715-52-1
List Price $106.00 · MPS/SIAM Member Price $74.20 · **Code MP04**

### A Mathematical View of Interior-Point Methods in Convex Optimization
James Renegar
2001 · viii + 117 pages · Softcover · ISBN 978-0-898715-02-6
List Price $47.00 · MPS/SIAM Member Price $32.90 · **Code MP03**

### Lectures on Modern Convex Optimization: *Analysis, Algorithms, and Engineering Applications*
Aharon Ben-Tal and Arkadi Nemirovski
2001 · xvi + 488 pages · Softcover · ISBN 978-0-898714-91-3
List Price $121.50 · MPS/SIAM Member Price $85.05 · **Code MP02**

### Trust-Region Methods
A. R. Conn, N. I. M. Gould, and Ph. L. Toint
2000 · xx + 959 pages · Hardcover · ISBN 978-0-898714-60-9
List Price $146.50 · MPS/SIAM Member Price $102.55 · **Code MP01**




***Also of Interest:***

### Assignment Problems
Rainer Burkard, Mauro Dell'Amico, and Silvano Martello
2009 · xx + 382 pages · Hardcover
ISBN 978-0-898716-63-4 · List Price $110.00
SIAM Member Price $77.00 · **Code OT106**

## YOU ARE INVITED TO CONTRIBUTE

The goal of the MPS-SIAM series is to publish a broad range of titles in the field of optimization and mathematical programming, characterized by the highest scientific quality.

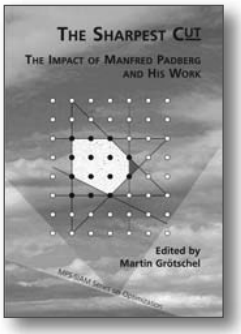
If you are interested in submitting a proposal or manuscript for publication in the series or would like additional information, please contact:

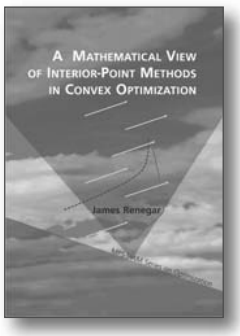
Philippe Toint
University of Namur
*philippe.toint@fundp.ac.be*

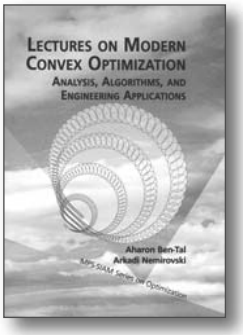
**OR**

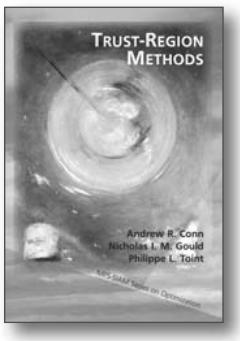
Sara J. Murphy
Series Acquisitions Editor
SIAM
*murphy@siam.org*

SIAM publishes quality books with practical implementation at prices affordable to individuals.




siam®

**Complete information about SIAM and its book program can be found at *www.siam.org/books*. See summaries, tables of contents, and order online at *www.siam.org/catalog*.**