On the Use of Intervals in Scientific Computing: What is the Best Transition from Linear to Quadratic Approximation?

Martine Ceberio[1], Vladik Kreinovich[1], and Lev Ginzburg[2,3]

[1] Department of Computer Science, University of Texas
El Paso, TX 79968, USA, {mceberio,vladik}@cs.utep.edu
[2] Department of Ecology and Evolution
State University of New York Stony Brook
NY 11794, USA, lev@ramas.com
[3] Applied Biomathematics, 100 N. Country Road
Setauket, NY 11733, USA

Abstract. In many problems from science and engineering, the measurements are reasonably accurate, so we can use linearization (= sensitivity analysis) to describe the effect of measurement errors on the result of data processing.

In many practical cases, the measurement accuracy is not so good, so, to get a good estimate of the resulting error, we need to take quadratic terms into consideration – i.e., in effect, approximate the original algorithm by a quadratic function. The problem of estimating the range of a quadratic function is NP-hard, so, in the general case, we can only hope for a good heuristic.

Traditional heuristic is similar to straightforward interval computations: we replace each operation with numbers with the corresponding operation of interval arithmetic (or of the arithmetic that takes partial probabilistic information into consideration). Alternatively, we can first diagonalize the quadratic matrix – and then apply the same approach to the result of diagonalization.

Which heuristic is better? We show that sometimes, the traditional heuristic is better; sometimes, the new approach is better; asymptotically, which heuristic is better depends on how fast, when sorted in decreasing order, the eigenvalues decrease.

1 Formulation of the Problem

Need for data processing and indirect measurements in scientific computing. In many areas of science and engineering, we are interested in the value of a physical quantity y that is difficult (or even impossible) to measure directly. Examples may include the amount of a pollutant in a given lake, the distance to a faraway star, etc.

To measure such quantities, we find auxiliary easier-to-measure quantities $x_1, \ldots, x_n$ that are related to y by a known algorithm $y = f(x_1, \ldots, x_n)$. In some cases, the relation between x_i and y is known exactly. In such cases, to estimate y, we measure x_i, and apply the algorithm f to the results $\widetilde{x}_1, \ldots, \widetilde{x}_n$ of measuring x_i. As a result, we get an estimate $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ for y.

In many other practical situations, we only know an approximate relation $y \approx \widetilde{f}(x_1, \ldots, x_n)$, with an upper bound ε_f on the accuracy of this approximation:

$$|\widetilde{f}(x_1, \ldots, x_n) - f(x_1, \ldots, x_n)| \le \varepsilon_f.$$

In such cases, to estimate y, we measure x_i, and apply the algorithm $\widetilde{f}$ to the results $\widetilde{x}_1, \ldots, \widetilde{x}_n$ of measuring x_i. As a result, we get an estimate $\widetilde{y} = \widetilde{f}(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ for y.

This indirect measurement (data processing) is one of the main reasons why computers were invented in the first place, and one of the main uses of computers is scientific computing.

Need for error estimation for indirect measurements in scientific computing. Measurements are never 100% accurate. The results $\widetilde{x}_i$ of direct measurements are, in general, different from the actual values x_i. Therefore, the estimate $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ is, in general, different from the actual (unknown) value $y = f(x_1, \ldots, x_n)$. What do we know about the error $\Delta y \stackrel{\text{def}}{=} \widetilde{y} - y$ of the indirect measurement?

Estimating errors of indirect measurements: formulation of the problem. In many cases, we know the upper bounds Δ_i on the measurement errors $\Delta x_i \stackrel{\text{def}}{=} \widetilde{x}_i - x_i$ of direct measurements. Once we know such an upper bound, we can guarantee that the actual value x_i lies in the interval $\mathbf{x}_i \stackrel{\text{def}}{=} [\widetilde{x}_i - \Delta_i, \widetilde{x}_i + \Delta_i]$. In this case, if we know the relation $y = f(x_1, \ldots, x_n)$ exactly, then the only information that we have about y is that y belongs to the range $[\underline{r}, \overline{r}] \stackrel{\text{def}}{=} f(\mathbf{x}_1, \ldots, \mathbf{x}_n)$.

In situations when, instead of knowing the exact relation $y = f(x_1, \ldots, x_n)$, we only know:

– the approximate relation $y \approx \widetilde{f}(x_1, \ldots, x_n)$ between x_i and y and
– we know the upper bound ε_f on the accuracy of approximating f by $\widetilde{f}$,

then we can guarantee that y belongs to the interval $[\underline{r} - \varepsilon_f, \overline{r} + \varepsilon_f]$, where $[\underline{r}, \overline{r}] \stackrel{\text{def}}{=} \widetilde{f}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is the range of a known algorithmic function $\widetilde{f}(x_1, \ldots, x_n)$ on the "box" $\mathbf{x}_1 \times \ldots \times \mathbf{x}_n$.

In both cases, to find the range of possible values of y, we must find the range $[\underline{r}, \overline{r}]$ of a known algorithmic function f (or $\widetilde{f}$) on the known box.

Comment. In some engineering situations, instead of knowing the guaranteed upper bounds Δ_i on the measurement errors, we only have *estimates* Δ_i of the upper bounds. In such situations, it is still desirable to compute the corresponding range for y – but we can no longer *absolutely* guarantee that the actual value y belong to the resulting range; we can only guarantee it *under the condition* that the estimates are correct.

Interval computations: a way to estimate errors of indirect measurements. Interval computations enable us to either compute the range a given algorithmic function f (or $\widetilde{f}$) on the given box exactly, or at least to provide an enclosure for this range. For the case when $n = 2$ and the function $f(x_1, x_2)$ is one of the standard arithmetic operations ($+$, $-$, multiplication, etc.), there are known explicit formulas for the range of f. For example,

$$[\underline{x}_1, \overline{x}_1] + [\underline{x}_2, \overline{x}_2] = [\underline{x}_1 + \underline{x}_2, \overline{x}_1 + \overline{x}_2].$$

These formulas form *interval arithmetic*; see, e.g., [3, 4, 7].

One way to compute the range for more complex functions f is to use straightforward ("naive") interval computations, i.e., replace each operation forming the algorithm f with the corresponding operation from interval arithmetic. This technique leads to an interval that is guaranteed to be an enclosure, i.e., to contain the desired range, but it is known that this interval contains *excess width*, i.e., is wider than the desired range [3, 4, 7]. How can we reduce this excess width?

When measurement errors are small, linearization works well. When the measurement errors Δx_i are relatively small, we can expand f into Taylor series in terms of Δx_i and ignore quadratic and higher terms – i.e., keep only linear terms. In a linear expression $f = a_0 + a_1 \cdot \Delta x_1 + \ldots + a_n \cdot \Delta x_n$, each variable $\Delta x_i \in [-\Delta_i, \Delta_i]$ occurs only once. It is known that for such *single-use expressions* (SUE), straightforward interval computations leads to the exact range; see, e.g., [2, 3].

Quadratic approximation is more difficult to analyze. In many real-life situations, measurement errors Δx_i are not so small, so we must also take into consideration terms that are quadratic in Δx_i. So, we must be able to estimate the range of a quadratic function

$$f = a_0 + \sum_{i=1}^{n} a_i \cdot \Delta x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \cdot \Delta x_i \cdot \Delta x_j,$$

or, equivalently,

$$f = a_0 + \sum_{i=1}^{n} a_i \cdot \Delta x_i + \sum_{i=1}^{n} a_{ii} \cdot (\Delta x_i)^2 + \sum_{i=1}^{n} \sum_{j \neq i} a_{ij} \cdot \Delta x_i \cdot \Delta x_j. \tag{1}$$

There exist methods for computing the exact range of such a function (see, e.g., [4]), but all such methods require 2^n steps – the number of steps which, even for a realistically large number of inputs $n \approx 10^2 - 10^3$, can be impossibly large. Since the problem of estimating range of a given quadratic function is, in general, NP-hard (see, e.g., [6, 9]), we cannot hope to get an algorithm that is always faster. So, for large n, we can only compute enclosures.

Two natural approaches to compute enclosure: which is better? One approach to computing the enclosure of a quadratic approximation function (1) is to use naive (straightforward) interval computations. As we have mentioned, in this approach, we often get excess width.

There is a particular case when we do not have any excess width – when the matrix $A = (a_{ij})_{i,j}$ is diagonal. In this case, f can be represented as a sum of the terms $a_i \cdot \Delta x_i + a_{ii} \cdot \Delta x_i^2$ corresponding to different variables, and each of these terms can be reformulated as a SUE expression $a_{ii} \cdot (\Delta x_i + a_i/(2a_{ii}))^2 + \mathrm{const}$ – thus making the whole expression SUE.

Every quadratic function can be represented in a similar diagonal form – as a linear combination of squares of eigenvectors. It therefore seems reasonable to first represent a quadratic function in this form, and only then apply straightforward interval computations.

A natural question is: which approach is better? If none of them is *always* better, then when is the first approach better and when is the second one better?

Beyond interval computations: towards joint use of probabilities and intervals in scientific computing. In many cases, in addition to the upper bounds on Δx_i, we have partial information on the probabilities of different values of $\Delta x \stackrel{\rm def}{=} (\Delta x_1, \ldots, \Delta x_n)$.

In particular, in some applications, we know that the input variables x_i are not truly independent and are in fact correlated. This knowledge about correlation is also usually represented in the probabilistic terms, as partial information about the probability distribution of Δx.

In all such cases, in addition to the interval range, we would like to compute the information about the probabilities of different values of y. There exist ways of extending interval arithmetic to such cases; see, e.g., [1]. We can therefore use both approaches in these cases as well.

What we are planning to do. In this paper, we show that which method is better depends on the eigenvalues of the matrix $B = (a_{ij} \cdot \Delta_i \cdot \Delta_j)_{i,j}$: on average, the eigenvector method is better if and only if the eigenvalues (when sorted in decreasing order) decrease fast enough.

2 Formalizing the Problem in Precise Terms

Simplifying the problem. Let us start by simplifying the above problem.

In the original formulation of the problem, we have parameters a_0, a_i, and a_{ij} that describe the function f and the parameters Δ_i that describe the accuracy of measuring each of n variables. We can reduce the number of parameters if we re-scale each of n variables in which a way that Δ_i becomes 1. Indeed, instead of the variables Δx_i, let us introduce the new variables $y_i \stackrel{\rm def}{=} \Delta x_i/\Delta_i$. For each of y_i, the interval of possible values is $[-1, 1]$. Substituting $\Delta x_i = \Delta_i \cdot y_i$ into the expression (1), we get the expression for f in terms of y_i:

$$f = b_0 + \sum_{i=1}^{n} b_i \cdot y_i + \sum_{i=1}^{n} b_{ii} \cdot y_i^2 + \sum_{i=1}^{n} \sum_{j \neq i} b_{ij} \cdot y_i \cdot y_j, \tag{2}$$

where $b_0 \stackrel{\text{def}}{=} a_0$, $b_i \stackrel{\text{def}}{=} a_i \cdot \mathit{\Delta}_i$, and $b_{ij} \stackrel{\text{def}}{=} a_{ij} \cdot \mathit{\Delta}_i \cdot \mathit{\Delta}_j$.

In the following text, we will therefore assume that $\mathit{\Delta}_i = 1$ and that the quadratic form has the form (2).

Explicit expressions for the results of the two compared methods. Let us explicitly describe the results of applying the two methods to the quadratic form (2).

If we directly apply straightforward interval computations to the original expression (2), then, since $y_i \in [-1, 1]$, we get the enclosure $f^{(0)} + \mathbf{f}^{(1)} + \mathbf{f}^{(2)}_{\text{orig}}$, where $f^{(0)} = b_0$, $\mathbf{f}^{(1)} = [-\sum\limits_{i=1}^{n} |b_i|, \sum\limits_{i=1}^{n} |b_i|]$, and

$$\mathbf{f}^{(2)}_{\text{orig}} = \sum_{i=1}^{n} (b_{ii} \cdot [0, 1]) + \sum_{i=1}^{n} \sum_{j \neq i} |b_{ij}| \cdot [-1, 1]. \tag{3}$$

Alternatively, we can represent the matrix $B = (b_{ij})_{i,j}$ in terms of its eigenvalues λ_k and the corresponding unit eigenvectors $e_k = (e_{k1}, \ldots, e_{kn})$, as

$$b_{ij} = \sum_{k=1}^{n} \lambda_k \cdot e_{ki} \cdot e_{kj}. \tag{4}$$

In this case, the original expression (2) takes the form

$$b_0 + \sum_{i=1}^{n} b_i \cdot y_i + \sum_{k=1}^{n} \lambda_k \cdot \left(\sum_{i=1}^{n} e_{ki} \cdot y_i \right)^2. \tag{5}$$

Since $y_i \in [-1, 1]$, we conclude that $\sum\limits_{i=1}^{n} e_{ki} \cdot y_i \in [-B_k, B_k]$, where $B_k \stackrel{\text{def}}{=} \sum\limits_{i=1}^{n} |e_{ki}|$. Therefore, $\left(\sum\limits_{i=1}^{n} e_{ki} \cdot y_i \right)^2 \in [0, B_k^2]$, and so, when applied to the expression (5), straightforward interval computations lead to the expression $f^{(0)} + \mathbf{f}^{(1)} + \mathbf{f}^{(2)}_{\text{new}}$, in which linear terms $f^{(0)}$ and $\mathbf{f}^{(1)}$ are the same, while

$$\mathbf{f}^{(2)}_{\text{new}} = \sum_{k=1}^{n} \lambda_k \cdot \left[0, \left(\sum_{i=1}^{n} |e_{ki}| \right)^2 \right]. \tag{6}$$

So, to decide which method is better, it is sufficient to consider only quadratic terms.

Example when the eigenvalue-related expression is better. If the matrix B has only one non-zero eigenvector $\lambda_1 \neq 0$, then the formula (5) takes a simplified form: $\lambda_1 \cdot (\sum\limits_{i=1}^{n} e_{1i} \cdot y_i)^2$. This is a SUE expression, so straightforward interval computations lead to the exact range.

For such matrices, the original expression (1) is not necessarily SUE, and may lead to excess width. For example, for a 2×2 matrix with $b_{ij} = 1$ for all i and j, the only non-zero eigenvalue is $\lambda_1 = 2$ (with eigenvector $(1,1)$). So, the new expression leads to the exact range $[0,4]$. On the other hand, if we apply straightforward interval computations to the original expression (2), then the resulting expression (3) leads to $[-2,4]$, i.e., to excess width.

Example when the original expression is better. For the identity matrix B, the original quadratic expression (2) leads to a SUE expression $\sum b_{ii} \cdot (\Delta x_i)^2$ for which straightforward interval computations lead to the exact range. For example, for $n = 1$, we get the range $[0,2]$.

On the other hand, if we select eigenvectors that are different from $(1,0)$ and $(0,1)$, we may get excess width. For example, if we choose $e_1 = (\sqrt{2}/2, \sqrt{2}/2)$ and $e_2 = (\sqrt{2}/2, -\sqrt{2}/2)$, then, for straightforward interval computations, the range of $\frac{\sqrt{2}}{2}\Delta x_1 + \frac{\sqrt{2}}{2}\Delta x_2$ is $[-\sqrt{2}, \sqrt{2}]$, hence the range of its square is $[0,2]$, and the range of the resulting quadratic expression is estimated as $[0,4]$.

How do we compare different approaches: randomization needed. The main difference between the two cases is in the eigenvalues of the matrix B: In the first example, we had only one non-zero eigenvalue, and the eigenvalue-related expression leads to better estimates. In the second example, we have equal eigenvalues, and the original expression is better. It is therefore natural to assume that which method is better depends on the eigenvalues λ_k of the matrix B.

We should not expect a result of the type "if we have certain λ_k, then the first method is always better" – which method is better depends also on the eigenvectors. For example, in the second case, if we select $(1,0)$ and $(0,1)$ as eigenvectors, then the eigenvalue-related expression also leads to the same optimal range estimate. In other words, for a given set of eigenvalues λ_k, we should not expect a result saying that one of the methods is better for *all* possible eigenvectors: for some eigenvectors the first methods will be better, for some others the second method will be better. In such a situation, it is reasonable to analyze which method is better *on average*, if we consider random eigenvectors.

Natural probability measure on the set of all eigenvectors. What is the natural probability measure on the set of all possible eigenvectors $e_1, \ldots, e_n$? In general, we have n mutually orthogonal unit vectors, i.e., an orthonormal base in the n-dimensional space. It is reasonable to assume that the probability distribution on the set of all such bases is rotation-invariant. This assumption uniquely determines the probability distribution; see, e.g., [5, 8].

Indeed, the first unit vector e_1 can be uniquely represented by its endpoint on a unit sphere. The only possible rotation-invariant distribution on a unit sphere is a uniform distribution. Once e_1 is fixed, e_2 can be any vector from a sphere in an $(n-1)$-dimensional space of all vectors orthogonal to e_1; the only rotation-invariant distribution on this sphere is also uniform, etc. So, in the resulting distribution, e_1 is selected from the uniform distribution on the unit sphere, e_2 from the uniform distribution on the unit sphere in the subspace of all vectors $\perp e_1$, etc.

3 Main Result

Theorem 1. *When $n \to \infty$, then asymptotically, the expected values are:*

$$E[\mathbf{f}^{(2)}_{\text{orig}}] \sim \left[-\sqrt{\frac{2}{\pi}} \cdot n \cdot \sqrt{\sum_{k=1}^{n} \lambda_k^2}, \sqrt{\frac{2}{\pi}} \cdot n \cdot \sqrt{\sum_{k=1}^{n} \lambda_k^2}\right]; \tag{7}$$

$$E[\mathbf{f}^{(2)}_{\text{new}}] \sim \left[\frac{2}{\pi} \cdot n \cdot \sum_{k:\lambda_k<0} \lambda_k, \frac{2}{\pi} \cdot n \cdot \sum_{k:\lambda_k>0} \lambda_k\right]. \tag{8}$$

Conclusions. If $\sum |\lambda_k| < \sqrt{\pi/2} \cdot \sqrt{\sum \lambda_k^2}$, then asymptotically, $E[\mathbf{f}^{(2)}_{\text{new}}] \subset E[\mathbf{f}^{(2)}_{\text{orig}}]$, so the eigenvector-based method is definitely better.

If $\sum |\lambda_k| < \sqrt{2\pi} \cdot \sqrt{\sum \lambda_k^2}$, then the interval $E[\mathbf{f}^{(2)}_{\text{new}}]$ is narrower than $E[\mathbf{f}^{(2)}_{\text{orig}}]$, so in this sense, the new method is also better.

Example. The spectrum λ_k often decreases according to the power law $\lambda_k \sim k^{-\alpha}$. In this case, $\sum |\lambda_k| \approx \int\limits_1^\infty x^{-\alpha}\,\mathrm{d}x = 1/(\alpha-1)$ and $\sum \lambda_k^2 \approx \int\limits_1^\infty x^{-2\alpha}\,\mathrm{d}x = 1/(2\alpha-1)$, so the above inequality turns into $(\alpha-1)^2 \ge (2/\pi)\cdot(2\alpha-1)$, which is equivalent to

$$\alpha \ge 1 + \frac{2}{\pi} + \sqrt{\left(1+\frac{2}{\pi}\right)\cdot\frac{2}{\pi}} \approx 2.7. \tag{9}$$

Hence, if the eigenvalues decrease fast ($\alpha \ge 2.7$), the new method is definitely better. For $\alpha \ge 1.6$, the new method leads to narrower intervals; otherwise, the traditional method leads, on average, to better estimates.

Proof of the Theorem. Before we start the proof, let us derive some auxiliary formulas. Since each vector e_k is a unit vector, we have $\sum\limits_i e_{ki}^2 = 1$. Due to rotation invariance, the expected value $E[e_{ki}^2]$ should not depend on i, hence $E[e_{ki}^2] = 1/n$. Similarly, from $\sum\limits_i e_{ki} \cdot e_{li} = 0$ and rotation invariance, we conclude that $E[e_{ki} \cdot e_{li}] = 0$.

For given k, l, and $i \ne j$, the value $E[e_{ki} \cdot e_{lj}]$ should not change under the transformation $x_i \to x_i$ and $x_j \to -x_j$, so $E[e_{ki} \cdot e_{lj}] = 0$.

To compute $E[\mathbf{f}^{(2)}_{\text{orig}}]$, we must find $E[b_{ii}]$ and $E[|b_{ij}|]$. By definition (4), for each i, $E[b_{ii}] = \sum\limits_k \lambda_k \cdot E[e^2_{ki}] = (1/n) \cdot \sum \lambda_k$, so the sum of n such terms is proportional to $\sum \lambda_k$.

For $i \neq j$, due to the central limit theorem, the distribution for b_{ij} (formula (4)) is asymptotically Gaussian, so asymptotically, $E[|b_{ij}|] \sim \sqrt{2/\pi} \cdot \sqrt{E[b^2_{ij}]}$. Here, $E[b^2_{ij}] = \sum\limits_k \sum\limits_l \lambda_k \cdot \lambda_l \cdot E[e_{ki} \cdot e_{kj} \cdot e_{li} \cdot e_{lj}]$. Due to symmetry, each $k \neq l$ term is 0, so $E[b^2_{ij}] = \sum\limits_k \lambda^2_k \cdot E[e^2_{ki} \cdot e^2_{kj}]$. Asymptotically, e_{ki} and e_{kj} are independent, so $E[e^2_{ki} \cdot e^2_{kj}] \sim E[e^2_{ki}] \cdot E[e^2_{kj}] = (1/n)^2$. Therefore, $E[b^2_{ij}] \sim (1/n)^2 \cdot \sum \lambda^2_k$, hence $E[|b_{ij}|] \sim \sqrt{2/n} \cdot (1/n) \cdot \sqrt{\sum \lambda^2_k}$. The sum of $n(n-1)$ such terms is $\sim \sqrt{2/\pi} \cdot n \cdot \sqrt{\lambda^2_k}$. The sum of the terms $E[b_{ii}]$ is asymptotically smaller, so when $n \to \infty$, we get the expression (7).

For the new expression, we must compute, for every k, the expected value of $E\left[\left(\sum\limits_i |e_{ki}|\right)^2\right] = \sum\limits_{i,j} E[|e_{ki}| \cdot |e_{kj}|]$. Asymptotically, e_{ki} and e_{kj} are independent, and $E[|e_{ki}|] \sim \sqrt{2/\pi} \cdot \sqrt{E[e^2_{ki}]} = \sqrt{2/\pi} \cdot (1/\sqrt{n})$. Thus, the sum of all the terms $i \neq j$ is $\sim n^2 \cdot (2/\pi) \cdot (1/n) = (2/\pi) \cdot n$. The terms with $i = j$ are asymptotically smaller, so we get the desired expression (8). □

Acknowledgments.

This work was supported in part by NASA grant NCC5-209, by the Air Force Office of Scientific Research grant F49620-00-1-0365, by NSF grants EAR-0112968, EAR-0225670, and EIA-0321328, and by the Army Research Laboratories grant DATM-05-02-C-0046. This work was partly performed during V. Kreinovich's visit to Brazil. This visit was sponsored by the Brazilian funding agency CT-INFO/CNPq.

The authors are thankful to the anonymous referees for valuable suggestions.

References

1. Ferson, S.: RAMAS Risk Calc 4.0, CRC Press, Boca Raton, Florida, 2002.
2. Hansen, E.: Sharpness in interval computations, Reliable Computing **3** (1997) 7–29.
3. Jaulin, L., Kieffer, M., Didrit, O., Walter, E.: Applied Interval Analysis: With Examples in Parameter and State Estimation, Robust Control and Robotics, Springer, London, 2001.
4. Kearfott, R.B.: Rigorous Global Search: Continuous Problems. Kluwer, Dordrecht, 1996.
5. Koltik, E., Dmitriev, V.G., Zheludeva, N.A., Kreinovich, V.: An optimal method for estimating a random error component, Investigations in Error Estimation, Proceedings of the Mendeleev Metrological Institute, Leningrad, 1986, 36–41 (in Russian).
6. Kreinovich, V., Lakeyev, A., Rohn, J., Kahl, P.: Computational Complexity and Feasibility of Data Processing and Interval Computations, Kluwer, Dordrecht, 1997.

7. Moore, R.E.: Methods and Applications of Interval Analysis. SIAM, Philadelphia, 1979.
8. Trejo, R., Kreinovich, V.: Error Estimations for Indirect Measurements: Randomized vs. Deterministic Algorithms For "Black-Box" Programs, In: Rajasekaran, S., Pardalos, P., Reif, J., Rolim, J., eds., Handbook on Randomized Computing, Kluwer, 2001, 673-729.
9. Vavasis, S.A.: Nonlinear Optimization: Complexity Issues, Oxford University Press, New York, 1991.