Optimal Approximation of Quadratic Interval Functions

Misha Koshelev[1] and Patrick Taillibert[2]

[1]1003 Robinson
El Paso, TX 79902, USA
email mkosh@cs.utep.edu

[2]Dassault Electronique
55 Quai Marcel Dassault
92214 Saint Cloud, France
email Patrick.Taillibert@dassault-elec.fr

Abstract

Measurements are never absolutely accurate; as a result, after each measurement, we do not get the *exact* value of the measured quantity; at best, we get an *interval* of its possible values.

For dynamically changing quantities x, the additional problem is that we cannot measure them continuously; we can only measure them at certain discrete moments of time $t_1, t_2, \ldots$. If we know that the value $x(t_j)$ at a moment t_j of the last measurement was in the interval $[x^-(t_j), x^+(t_j)]$, and if we know the upper bound D on the rate with which x changes (i.e., on the absolute value $|\dot{x}(t)|$ of the derivative $\dot{x}(t)$), then, for any given moment of time t, we can conclude that $x(t)$ belongs to the interval $[x^-(t_j) - D \cdot (t - t_j), x^+(t_j) + D \cdot (t - t_j)]$. This interval changes *linearly* with time, an is, therefore, called a *linear interval function*.

When we process these intervals, we get an expression that is quadratic and higher order w.r.t. time t. Such "quadratic" intervals are difficult to process and therefore, it is necessary to approximate them by linear ones.

In this paper, we describe an algorithm that gives the optimal approximation of quadratic interval functions by linear ones.

1 Introduction: intervals, linear and quadratic interval functions, and why it is necessary to approximate

The need for indirect measurements. In many real-life problems, we are interested in the values of a quantity y that is difficult or impossible to measure directly. For example:

- in *astrophysics*, we cannot directly measure the mass or the temperature of the star;
- in *geophysics*, we cannot directly measure the amount of oil in a given area, etc.

In all these situations, we measure some related quantities $x_1, \ldots, x_n$ that *can* be measured directly, and then use the known relationship $y = f(x_1, \ldots, x_n)$ between x_i and y and the results $\tilde{x}_i$ of measuring x_i to estimate y as $\tilde{y} = f(\tilde{x}_1, \ldots, \tilde{x}_n)$. For example:

- to measure the temperature of the star, we measure its brightness $x_1, \ldots, x_n$ on different wavelengths, and then try to fit the resulting spectrum into a black-body radiation spectrum corresponding to the temperature y;
- to estimate the amount of oil y in a given area, we measure the geophysical characteristics of different parts of it, and use known equations to estimate y.

In contrast to a direct measurement of x_i, this two-stage process of direct measurements followed by data processing is called *indirect measurement.*

In space sciences, and especially in their applications to environmental and earth sciences, indirect measurements are the main method of measurement.

The main problem: estimating accuracy of the result of data processing. After the data processing is over, the main problem is: how accurate is the result $\tilde{y}$ of data processing?

For example, suppose that we estimate the amount of oil in a given area, and we got $\tilde{y} = 100$ mln ton. Our further actions will depend on the accuracy of this result:

- if the accuracy is high, e.g., if the actual amount of oil is 100 ± 1, then we will probably start drilling;
- however, if the accuracy is low, and the actual value if, say, 100 ± 100, then it may make more sense to undertake further, more accurate, measurements.

The inaccuracy in y comes from two sources:

- First, the model $y = f(x_1, \ldots, x_n)$ that is used in data processing is often inaccurate. This inaccuracy is the easiest to take into consideration: if we know the upper bound Δ_m on the error $y - f(x_1, \ldots, x_n)$ of this model, and if we know the values of x_i, then we can conclude that the actual value of y belongs to the interval $[\tilde{y} - \Delta_m, \tilde{y} + \Delta_m]$, where $\tilde{y} = f(x_1, \ldots, x_n)$.

- Second, the measurement results $\tilde{x}_i$ are also inaccurate and therefore, may differ from the actual values of the measured quantities x_i. This error is much more difficult to estimate, and this is what we will be doing in the present paper.

Intervals. Measurements are never 100% precise. Thus, if, as the result of measuring a certain quantity x_i, we get a measurement result $\tilde{x}_i$, it does not necessarily mean that the actual value x_i of this quantity is exactly equal to $\tilde{x}_i$.

The manufacturer of the measuring instrument usually supplies it with the upper bound Δ_i for the measurement error $\Delta x_i = \tilde{x}_i - x_i$; in order words, the manufacturer guarantees that $|\Delta x_i| \leq \Delta_i$.

In this case, if we have measured a quantity x_i and the measurement result is $\tilde{x}_i$, then the only information that we have about the actual value is that this actual value cannot differ from $\tilde{x}_i$ by more than Δ_i, i.e., that this actual value must be within the *interval* $[\tilde{x}_i - \Delta_i, \tilde{x}_i + \Delta_i]$.

If we know these intervals, then the set Y of possible values of $y = f(x_1, \ldots, x_n)$ can be described as follows:

$$Y = \{f(x_1, \ldots, x_n) \,|\, x_1 \in [x_1^-, x_1^+], \ldots, x_n \in [x_n^-, x_n^+]\}. \tag{1}$$

Comments.

- Sometimes, in addition to the upper bound for the error, we know the *probabilities* of different error values. However, in many real-life cases, we do not know these probabilities, and the upper bound Δ_i is the only information about the measurement error Δx_i that we have.

- Computations that take this interval uncertainty into consideration are called *interval computations* (see, e.g., [1, 2]).

Linear interval functions. In some cases, the values x_i do not change with time. In these cases, the only inaccuracy is caused by the inaccuracy of the direct measurement.

However, in many real-life situations, the measured values change with time. For such dynamically changing quantities x_i, there is an additional source of uncertainty: namely, we cannot measure a quantity continuously; we can only measure it at certain discrete moments of time $t_1, t_2, \ldots$ Hence:

- If we are interested in the value of $x_i(t_j)$ at one of the moments of time t_j in which we have actually measured x_i, then the inaccuracy of the direct measurement is the only source of the measurement error, and we get the interval $[x_i^-(t_j), x_i^+(t_j)]$ of possible values of $x_i(t_j)$.

- However, if we are interested in the value $x_i(t)$ of the quantity x_i at a moment of time t in which no direct measurement of x_i was performed, then we must use one of the measured values, and get an additional error component caused by the change of x_i between the moment t_j of the last measurement and the desired moment t. There are two methods to take this additional error into consideration:

  - If we know the upper bound D_i on the rate with which x changes (i.e., on the absolute value $|\dot{x}_i(t)|$ of the derivative $\dot{x}_i(t)$), then, for any given moment of time t, we can conclude that $x_i(t_j) - D_i \cdot (t - t_j) \le x_i(t) \le x_i(t_j) + D_i \cdot (t - t_j)$. Hence, from the fact that $x_i^-(t_j) \le x_i(t_j) \le x^+(t_j)$, we conclude that the value $x_i(t)$ belongs to the interval $[x_i^-(t), x_i^+(t)]$, where $x_i^-(t) = x_i^-(t_j) - D_i \cdot (t - t_j)$ and $x_i^+(t) = x_i^+(t_j) + D_i \cdot (t - t_j)$.
  - In some cases, simultaneously with measuring the values $x_i(t_j)$, we can directly measure the rate $\dot{x}_i(t_j)$ with which these values change; e.g.,

    * simultaneously with measuring coordinates $x_i(t)$ of a robot, we can measure its velocity $\dot{x}_i(t)$;
    * simultaneously with measuring the velocity $x_i(t)$, we can measure the acceleration $\dot{x}_i(t)$; etc.

    In these case, if we know the interval of possible values $[\dot{x}_i^-(t_j), \dot{x}_i^+(t_j)]$ of the rate, and we know that the upper bound S_i on the absolute value of the second time derivative, we can conclude that:

    * for all $t \in [t_j, t_{j+1}]$, the values $\dot{x}_i(t_j)$ belong to the intervals $[v^-, v^+]$, where $v^- = \dot{x}_i^-(t_j) - S_i \cdot (t_{j+1} - t_j)$ and $v^+ = \dot{x}_i^+(t_j) + S_i \cdot (t_{j+1} - t_j)$, and therefore, that
    * for each of these moments of time, the actual values of $x_i(t)$ belong to the interval $[x^-(t), x^+(t)]$, where $x^-(t) = x_i^-(t_j) - v^- \cdot (t - t_j)$ and $x^+(t) = x_i^+(t_j) + v^+ \cdot (t - t_j)$.

In general, the only information that we have about the actual value of $x(t)$ is that $x(t) \in \mathbf{x}(t) = [x^-(t), x^+(t)]$. In other words, instead of a single *interval*, we have an *interval function* that assigns an interval $\mathbf{x}(t)$ to each moment of time t.

In our cases, the endpoints $x^-(t)$ and $x^+(t)$ of this interval are linear functions and therefore, it is natural to call this interval function *linear interval function*.

For linear data processing, linear interval functions remain linear. If the algorithm $f(x_1, \ldots, x_n)$ is a linear function, then, as a result of applying this algorithm to the interval linear functions, we get an estimate (1) that is either a linear or a piece-wise linear function of time t.

In [3, 4, 5, 6], such linear interval functions were effectively used in electrical and electronic engineering analysis of circuits with linear elements.

For non-linear data processing, we need an approximation. In most real-life situations, however, data processing algorithms are non-linear. As a result, if we start with the intervals for $x_i(t)$ that are linear in time t, we end up with intervals for $y(t)$ whose dependence on t is much more complicated. For example:

- if we *add* or *subtract* two intervals that are linear in time, we still get the result that is linear in time; but
- if we *multiply* two intervals that are linear in time, then their endpoints also get multiplied, and, as a result, we get an interval function $[y^-(t), y^+(t)]$ in which the endpoints are *quadratic* functions of time (i.e., we get *quadratic interval functions*).

If we multiply more, we get cubic, quartic, etc. functions; see, e.g., [6]. In principle, there is nothing wrong with this complexity, except for the fact that while a linear interval function requires only 4 numbers to store (2 coefficients of the lower endpoint and 2 coefficients of the upper endpoint), quadratic functions require 6 coefficients, cubic functions require 8, etc. The more complicated the function becomes, the more memory we need to store these coefficients, and the longer it takes to process these functions.

Thus, since we are often limited both in processing time and in memory (especially if the processing is done in an on-board computer), we must *approximate* the given complicated interval function by a simpler one.

In other words, if we have an interval function $\mathbf{y}(t) = [y^-(t), y^+(t)]$ that is known to contain the actual value $y(t)$, we want to be able to find a simpler interval function $\mathbf{z}(t) = [z^-(t), z^+(t)]$ that for each t, contains the entire interval $\mathbf{y}(t)$ and is, thus, guaranteed to contain the actual value $y(t)$.

Of course, this approximation comes at a trade-off: we simplify the expression, but we make the interval wider (and therefore, lose some information). Therefore, the narrower the approximating interval function, the better.

What we are planning to do. The simplest approximation problem of this type is the problem of approximating a quadratic interval function by linear ones. In this paper, we will present an optimal solution to this problem.

2 When is the approximation optimal? Mathematical formulation of the problem

Definition 1. *Let an interval* $[t^-, t^+]$ *be fixed.*

- *By an interval function, we mean a mapping* $\mathbf{x}$ *that puts into correspondence to each number* $t \in [t^-, t^+]$ *an interval* $\mathbf{x}(t) = [x^-(t), x^+(t)]$.
- *If both functions* $x^-(t)$ *and* $x^+(t)$ *are linear functions of* t *(i.e., if* $x^\pm(t) = x_0^\pm + x_1^\pm \cdot t$*), we say that the interval function is linear.*
- *If both functions* $x^-(t)$ *and* $x^+(t)$ *are quadratic functions of* t *(i.e., if* $x^\pm(t) = x_0^\pm + x_1^\pm \cdot t + x_2^\pm \cdot t^2$*), we say that the interval function is quadratic.*
- *We say that a linear interval function* $\mathbf{z}(t) = [z^-(t), z^+(t)]$ *approximates a quadratic interval function* $\mathbf{y}(t) = [y^-(t), y^+(t)]$ *if* $\mathbf{y}(t) \subseteq \mathbf{z}(t)$ *for all* t.

Comment. The narrower the intervals, the better. So, our goal is to minimize the worse-case width of the approximating interval, i.e., the value $W(\mathbf{z}) = \max_t(z^+(t) - z^-(t))$.

We will see that in some cases, for a given quadratic interval function $\mathbf{y}(t)$, there are several approximating linear interval functions $\mathbf{z}$ with the same value of $W(\mathbf{z})$. If two different approximating functions have the same worst-case widths, then it is reasonable to choose the one for which the best-case width $w(\mathbf{z}) = \min_t(z^+(t) - z^-(t))$ is the smallest. Thus, we arrive at the following definition:

Definition 2.

- *For every interval function* $\mathbf{z}(t)$:
  - *by its worst-case width, we mean a value*

$$W(\mathbf{z}) = \max_{t \in [t^-, t^+]} (z^+(t) - z^-(t)).$$

  - *by its best-case width, we mean a value*

$$w(\mathbf{z}) = \min_{t \in [t^-, t^+]} (z^+(t) - z^-(t)).$$

- *Let a quadratic interval function* $\mathbf{y}$ *be fixed. We say that a linear function* $\mathbf{z}$ *is an optimal approximation of* $\mathbf{y}$ *if the following conditions are satisfied:*
  - *first,* $\mathbf{z}$ *is an approximation of* $\mathbf{y}$.
  - *second, among all linear approximations to* $\mathbf{y}$*, the function* $\mathbf{z}$ *has the smallest value of the worst-case width* $W(\mathbf{z})$;

– *third, if there exist several linear approximations* $\mathbf{u}$ *to* $\mathbf{y}$*, with the same smallest value of the worst-case width* $W(\mathbf{u})$*, the function* $\mathbf{z}$ *has the smallest value of the best-case width* $w(\mathbf{z})$*.*

What we are planning to do. In this paper, we will describe the optimal linear approximation to an arbitrary quadratic interval function.

The resulting formulas will depend on the signs of the coefficients (y_2^- and y_2^+) at the quadratic term t^2. In each of the resulting four cases, we will describe the geometric idea that leads to the optimal approximation, and give the explicit formulas for this approximation.

3 Case 1: $y_2^- \le 0$, $y_2^+ \ge 0$

Derivation of the solution. In this case, both functions $y^-(x)$ and $y^+(x)$ are "pointing inward".

If $\mathbf{z}(t)$ is a linear interval approximation to the given quadratic interval function, then, from $\mathbf{y}(t) \subseteq \mathbf{z}(t)$, it follows, in particular, that $z^-(t) \le y^-(t)$ for all $t \in [t^-, t^+]$; therefore, $z^-(t^-) \le y^-(t^-)$ and $z^-(t^+) \le y^-(t^+)$.

Since $y_2^- \le 0$, the quadratic function $y^-(t)$ is concave and hence, from $z^-(t^-) \le y^-(t^-)$ and $z^-(t^+) \le y^-(t^+)$, i.e., from the fact that the line $z^-(t)$ lies below the two endpoints of the graph of $y^-(t)$, it automatically follows that the entire graph of $y^-(t)$ is above the line $z^-(t)$. So, it is sufficient to guarantee that at the endpoints, the values $z^-(t^\pm)$ do not exceed the corresponding values of $y^-(t^\pm)$.

For fixed $z^+(t)$, the resulting intervals are the narrowest when the line $z^-(t)$ is at the highest possible location. Thus, to minimize the widths of the intervals, we must move both points $z^-(t^\pm)$ up as much as possible. The highest possible location for $z^-(t^-)$ is $y^-(t^-)$, and the highest possible location for $z^-(t^+)$ is $y^-(t^+)$. Thus, $y^-(t)$ is a a straight line going through $y^-(t^\pm)$, i.e., a *secant.*

Similarly, $z^+(t)$ is a secant of $y^+(t)$.

Solution. In the optimal approximation, the function $z^-(t)$ is the *secant* of $y^-(t)$, i.e., a straight line whose endpoints are the endpoints of the quadratic function $y^-(t)$ on this interval. Similarly, $z^+(t)$ is a secant of $y^+(t)$, i.e.,

$$z^-(t) = y^-(t^-) + \frac{y^-(t^+) - y^-(t^-)}{t^+ - t^-} \cdot (t - t^-), \tag{2}$$

$$z^-(t) = y^+(t^-) + \frac{y^+(t^+) - y^+(t^-)}{t^+ - t^-} \cdot (t - t^-). \tag{3}$$

4 Case 2: $y_2^- \le 0$, $y_2^+ < 0$

Derivation of the solution. The arguments given for Case 1 show that in this case, the function $z^-(t)$ is still the secant. With $z^-(t)$ fixed, it is sufficient to find the upper function $z^+(t)$ for which the resulting approximation is optimal.

Let us first guarantee that the worst-case width is indeed the smallest possible. If this width is equal to W_0, this means that $z^+(t) \le z^-(t) + W_0$, and therefore, that $y^+(t) \le z^+(t) \le z^-(t) + W_0$. Thus, to guarantee that W_0 takes the smallest possible values, we must choose W_0 as the smallest possible value for which $y^+(t) \le z^-(t) + W_0$ for all $t \in [t^-, t^+]$. In other words, as W_0, we take the maximum of the function $y^+(t) - z^-(t)$ on the interval $[t^-, t^+]$.

The maximum of the concave quadratic function $y^+(t) - z^-(t) = (y_0^+ - z_0^-) + (y_1^+ - z_1^-) \cdot t + y_2^- \cdot t^2$ is attained at the point where its derivative is equal to 0, i.e., at the point

$$t_m = -(y_1^+ - z_1^-)/(2y_2^+). \tag{4}$$

If this maximum is attained at the internal point t_m of this interval, then at t_m, the line $z^-(t) + W_0$ is a tangent to $y^+(t)$, and therefore, there is no way to find a lower line without increasing the worst-case width $W(\mathbf{z})$. So, in this case, $z^+(t) = z^-(t) + W_0$.

If the maximum is attained in one of the endpoints, e.g., at t^+, then, we can, keeping the straight line $z^+(t)$ at the point $(t^+, y^+(t^+))$, lower its other end and still get the same worst-case width. The lowest value of the best-case width is attained when we lower the other end to the lowest possible position in which it is still above $y^+(t)$, i.e., to the position of a *tangent* to $y^+(t)$.

Solution. For the optimal approximation, the lower line $z^-(t)$ is a secant (2). To determine the upper line $z^+(t)$, we apply the formula (4) to compute the value t_m. Then:

- If $t_m \in [t^-, t^+]$, we take $z^+(t) = z^-(t) + (y^+(t_m) - z^-(t_m))$.
- If $t_m > t^+$, then z^+ is the tangent to y^+ at t^+:

$$z^+(t) = y^+(t^+) + (y_1^+ + 2y_2^+ \cdot t^+)(t - t^+). \tag{5}$$

- If $t_m < t^-$, then z^+ is the tangent to y^+ at t^-:

$$z^+(t) = y^+(t^-) + (y_1^+ + 2y_2^+ \cdot t^-)(t - t^-). \tag{6}$$

5 Case 3: $y_2^- > 0$, $y_2^+ \ge 0$

This case is similar to case 2, so we can immediately give a solution:

Solution. For the optimal approximation, the upper line $z^+(t)$ is a secant (3). To determine the lower line $z^-(t)$, we compute the value $t_m = -(y_1^- - z_1^+)/(2y_2^-)$. Then:

- If $t_m \in [t^-, t^+]$, we take $z^-(t) = z^+(t) - (z^+(t_m) - y^-(t_m))$.

- If $t_m > t^+$, then z^- is the tangent to y^- at t^+:

$$z^-(t) = y^-(t^+) + (y_1^- + 2y_2^- t^+)(t - t^+). \tag{7}$$

- If $t_m < t^-$, then z^- is the tangent to y^- at t^-:

$$z^-(t) = y^-(t^-) + (y_1^- + 2y_2^- t^-)(t - t^-). \tag{8}$$

6 Case 4: $y_2^- > 0$, $y_2^+ < 0$

Derivation of the solution. In this case, both functions $y^-(x)$ and $y^+(x)$ are "pointing outward".

As a first step of constructing the optimal linear approximation $\mathbf{z}(t)$, let us first make sure that we have the smallest possible value of the worst-case width. For every $t \in [t^-, t^+]$, from $\mathbf{y}(t) \subseteq \mathbf{z}(t)$, we can conclude that the width of $\mathbf{z}(t)$ is at least as large as the width of the interval $\mathbf{y}(t)$. Thus, the worst-case width $W(\mathbf{z})$ of the approximating linear function $\mathbf{z}(t)$ cannot be smaller than the worst-case width $W(\mathbf{y})$ of the original (quadratic) interval function $\mathbf{y}$. Since we are minimizing $W(\mathbf{z})$, it is therefore desirable to choose $\mathbf{z}$ in such a way that its worst-case width is *exactly equal* to $W(\mathbf{y})$.

The worst-case width $W(\mathbf{y})$ is a maximum of the quadratic width function $y^+(t) - y^-(t)$ on the interval $[t^-, t^+]$. By differentiating this difference, one can easily get an explicit expression for this maximum point t_M (see below).

If this maximum point t_M is inside the interval $[t^-, t^+]$, then at this point t_M, both approximating lines $z^-(t)$ and $z^+(t)$ must be tangent to the corresponding functions $y^-(t)$ and $y^+(t)$, because otherwise, in at least one of the directions, the width will increase.

If this maximum t_M is attained at one the endpoints, e.g., for t^+, then we must have $z^-(t^+) = y^-(t^+)$ and $z^+(t^+) = y^+(t^+)$. In this case, to guarantee the smallest possible *best-case* width, we must place $z^+(t)$ as low as possible (i.e., along the tangent to y^+), and $z^-(t)$ as high as possible, i.e., similarly, along the tangent to $y^-(t)$.

As a result, we arrive at the following formulas:

Solution. Compute $t_M = -(y_1^+ - y_1^-)/[2(y_2^+ - y_2^-)]$. Then:

- If $t_M \in [t^-, t^+]$, then z^- is the tangent to y^- at t_M, and z^+ is the tangent to y^+ at t_M: $z^-(t) = y^-(t_M) + (y_1^- + 2y_2^- \cdot t_M)(t - t_M)$; $z^+(t) = y^+(t_M) + (y_1^+ + 2y_2^+ \cdot t_M)(t - t_M)$.

- If $t_M > t^+$, then z^- is the tangent (7) to y^- at t_+, and z^+ is the tangent (5) to y^+ at t_+;

- If $t_M < t^-$, then z^- is the tangent (8) to y^- at t_- and z^+ is the tangent (6) to y^+ at t_-.

Acknowledgments. This work was partially supported by the NASA Pan American Center for Environmental and Earth Studies (PACES). The authors are thankful to Ann Gates, Vladik Kreinovich, Luc Longpré, and Scott Starks for their help and encouragement.

References

[1] R. B. Kearfott, *Rigorous global search: continuous problems*, Kluwer, Dordrecht, 1996.

[2] R. B. Kearfott and V. Kreinovich (eds.), *Applications of Interval Computations*, Kluwer, Dordrecht, 1996.

[3] O. Lhomme, "Consistency Techniques for Numeric CSPs," In: *Proceedings of IJCAI'93, the Thirteenth International Joint Conference on Artificial Intelligence*, 1993.

[4] O. Lhomme, A. Gottlieb, M. Rueher, and P. Taillibert, "Boosting the Interval Narrowing Algorithm," In: *Proceedings of JICSLP'96, Joint International Conference and Symposium on Logic Programming, Bonn, Germany, September 2–6*, 1996.

[5] E. Loiez and P. Taillibert, "Analog Systems Diagnosis: Modeling with Temporal Bands," In: *Proceedings of CESA'96, IEEE International Conference on Computations in Engineering and Simulations Applications, Lille, France, July 9–12, 1996.*

[6] E. Loiez and P. Taillibert, "Polynomial Temporal Band Sequences for Analog Diagnosis", In: *Proceedings of the 7th International Workshop on Principles of Diagnosis DX'96, Val Morin, Quebec, Canada, October 13-16, 1996.*