Solutions to Test 1 for Interval Class, Fall 2023

Problem 1. Why do we need data processing?

Answer. We want to predict the future state of the world, we want to know how to change this state. For this, we need to know the current state of the world. In general, the state of the world is characterized by numerical values of different quantities. Some of these quantities y we can measure directly, others we can only measure indirectly, by measuring easier-to-measure quantities $x_1, \ldots, x_n$ which are related to the desired quantity y by a known algorithm $y = f(x_1, \ldots, x_n)$. Using this algorithm to transform the results of measuring x_i into an estimate for y is one example of data processing.

Another example is using the known prediction algorithm $y = f(x_1, \ldots, x_n)$ to predict the future value y of some quantity based on the available information

$$x_1, \ldots, x_n.$$

Problem 2. Why do we need interval computations?

Answer. Measurements are never 100% accurate. The measurement result $\widetilde{x}$ is, in general, different from the actual (unknown) value of the corresponding quantity. There is, in general, a non-zero measurement error $\Delta x = \widetilde{x} - x$.

In many practical situations, the only information that we have about the measurement error is the upper bound Δ on its absolute value: $|\Delta x| \le \Delta$. In this case, once we know the measurement result $\widetilde{x}$, the only information that we have about the actual value x is that this value is located somewhere in the interval $[\widetilde{x} - \Delta, \widetilde{x} + \Delta]$.

Since the values $\widetilde{x}_i$ are, in general, different from the actual values x_i, the result $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ of data processing is, in general, different from the value $y = f(x_1, \ldots, x_n)$ that we would have gotten if we knew the actual values x_i.

In many practical situations, it is desirable to know how big can the difference $\Delta y = \widetilde{y} - y$ be, i.e., what is the range of possible values of $y = f(x_1, \ldots, x_n)$ when each x_i is in the corresponding interval $[\widetilde{x}_i - \Delta_i, \widetilde{x}_i + \Delta_i]$. Computing this range based on the known intervals for x_i is called *interval computation.*

Problem 3. What is the main problem of interval computations?

Answer. Given:

- algorithm $y = f(x_1, \ldots, x_n)$ that transforms n real numbers x_i into a real number, and
- n intervals $[\underline{x}_i, \overline{x}_i]$, $i = 1, \ldots, n$.

We want to find the range

$$[\underline{y}, \overline{y}] = \{f(x_1, \ldots, x_n) : x_i \in [\underline{x}_i, \overline{x}_i] \text{ for all } i\}.$$

Problem 4. Use calculus to find the range of a function $f(x) = (x-1)\cdot(2-x)$ on the interval $[1.2, 1.6]$. Explain the main algorithm of using calculus to find the range.

Solution. We consider all possible combinations $(x_1, \ldots, x_n)$ for which, for each i, we have:

- either $x_i = \underline{x}_i$,
- or $x_i = \overline{x}_i$,
- or $\dfrac{\partial f}{\partial x_i} = 0$ and $x_i \in [\underline{x}_i, \overline{x}_i]$.

For each such combination, we compute the value $f(x_1, \ldots, x_n)$. The smallest of these values of the function is $\underline{y}$, the largest is $\overline{y}$.

Maximum and minimum are attained either at the endpoints or where the derivative is equal to 0. Here,

$$f'(x) = 2 - x - (x-1) = 2 - x - x + 1 = 2x - 3,$$

so $f'(x) = 0$ when $2x - 3 = 0$, i.e., when $x = 1.5$. So, to find the endpoints $\underline{y}$ and $\overline{y}$ of the desired range, we must find the values $f(x)$ for $x = 1.2$, $x = 1.6$, and $x = 1.5$. We have:

- $f(1.2) = (1.2-1)\cdot(2-1.2) = 0.2 \cdot 0.8 = 0.16$;
- $f(1.6) = (1.6-1)\cdot(2-1.6) = 0.6 \cdot 0.4 = 0.24$;
- $f(1.5) = (1.5-1)\cdot(2-1.5) = 0.5 \cdot 0.5 = 0.25$.

The smallest of these three values in 0.16, the largest is 0.25, so the range is

$$[0.16, 0.25].$$

Problem 5. Use calculus to find the range of a function

$$f(x_1, x_2) = (x_1 - 1) \cdot (2 - x_2)$$

when x_1 is in the interval $[0.8, 1.2]$ and x_2 is in the interval $[-1.2, -0.8]$. Why cannot we use calculus to find the range for any number of inputs?

Solution. For each of the variables x_i, we need to consider cases when $x_i = \underline{x}_i$, when $x_i = \overline{x}_i$, and when $\dfrac{\partial f}{\partial x_i} = 0$.

- Here, $\dfrac{\partial f}{\partial x_1} = 2 - x_2$, so this derivative is equal to 0 when $x_2 = 2$, which is outside the range $[-1.2, -0.8]$ of x_2. So, we cannot have $\dfrac{\partial f}{\partial x_1} = 0$.
- Similarly, $\dfrac{\partial f}{\partial x_2} = -(x_1 - 1) = 1 - x_1$, so this derivative is equal to 0 when $x_1 = 1$, which is inside the range for x_1.

Let us first consider cases when $x_1 = 0.8$.

- If $x_2 = -1.2$, then
$$f(x_1, x_2) = (0.8 - 1) \cdot (2 - (-1.2)) = (-0.2) \cdot 3.2 = -0.64.$$
- If $x_2 = -0.8$, then
$$f(x_1, x_2) = (0.8 - 1) \cdot (2 - (-0.8)) = (-0.2) \cdot 2.8 = -0.56.$$
- If $\dfrac{\partial f}{\partial x_2} = 0$, then $x_1 = 1$, so we cannot have $x_1 = 0.8$ – this case is impossible.

If $x_1 = 1.2$, then:

- If $x_2 = -1.2$, then
$$f(x_1, x_2) = (1.2 - 1) \cdot (2 - (-1.2)) = 0.2 \cdot 3.2 = 0.64.$$
- If $x_2 = -0.8$, then
$$f(x_1, x_2) = (1.2 - 1) \cdot (2 - (-0.8)) = 0.2 \cdot 2.8 = 0.56.$$
- If $\dfrac{\partial f}{\partial x_2} = 0$, then $x_1 = 1$, so we cannot have $x_1 = 1.2$ – this case is impossible.

Finally, the case when $\dfrac{\partial f}{\partial x_1} = 0$ is not possible. These values can be described by the following table:

	$x_1 = 0.8$	$x_1 = 1.2$	$\dfrac{\partial f}{\partial x_1} = 0$
$x_2 = -1.2$	-0.64	0.64	X
$x_2 = -0.8$	-0.56	0.56	X
$\dfrac{\partial f}{\partial x_2} = 0$	X	X	X

The smallest of the values is -0.64, the largest is 0.64, so the range is $[-0.64, 0.64]$.

Why cannot we use calculus to find the range for any number of inputs? Because even if for each of the variables, we have only 2 options, for n variables, we need to consider 2^n combinations $(x_1, \ldots, x_n)$. For $n = 1000$, testing all these combinations would take longer than the lifetime of the universe.

Problem 6–7. Linearize the expression from Problem 4 around the interval's midpoint. Use the linearized expression to find the approximate value of the range of the original function, both with the actual derivative and with the result of numerical differentiation. Write down the general formulas for both linearization methods; in each formula, explain what each variable means.

Answer. In general, in addition to the data processing algorithm $f(x_1, \ldots, x_n)$, we are given:

- either the results $\widetilde{x}_i$ of measuring x_i and the upper bounds Δ_i on the absolute value of the corresponding measurement error,
- or the intervals $[\underline{x}_i, \overline{x}_i]$.

If we are given the intervals, then we need to compute the values

$$\widetilde{x}_i = \frac{\underline{x}_i + \overline{x}_i}{2} \text{ and } \Delta_i = \frac{\overline{x}_i - \underline{x}_i}{2}.$$

In both methods, the range is estimated as $[\widetilde{y} - \Delta, \widetilde{y} + \Delta]$. Here $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$.

- If we use actual derivative, then the value Δ is estimated as $\Delta = \sum\limits_{i=1}^{n} |c_i| \cdot \Delta_i$, where
$$c_i = \frac{\partial f}{\partial x_i}(\widetilde{x}_1, \ldots, \widetilde{x}_n).$$
- If we use numerical differentiation, then the value Δ is estimated as
$$\Delta = \sum_{i=1}^{n} |f(\widetilde{x}_1, \ldots, \widetilde{x}_{i-1}, \widetilde{x}_i + \Delta_i, \widetilde{x}_{i+1}, \ldots, \widetilde{x}_n) - \widetilde{y}|.$$

In this particular problem, the midpoint is $\widetilde{x}_1 = \dfrac{1.2 + 1.6}{2} = 1.4$ and the half-width is $\Delta_1 = \dfrac{1.6 - 1.2}{2} = 0.2$. So

$$\widetilde{y} = f(\widetilde{x}_1) = f(1.4) = (1.4 - 1) \cdot (2 - 1.4) = 0.4 \cdot 0.6 = 0.24.$$

If we use the actual derivative, then we get

$$c_1 = f'(\widetilde{x}_1) = 3 - 2 \cdot 1.4 = 3 - 2.8 = 0.2.$$

Thus,

$$\Delta = |c_1| \cdot \Delta_1 = |0.2| \cdot 0.2 = 0.2 \cdot 0.2 = 0.04.$$

Thus, the range is equal to

$$[\widetilde{y} - \Delta, \widetilde{y} + \Delta] = [0.24 - 0.04, 0.24 + 0.04] = [0.20, 0.28].$$

If we use numerical differentiation, then $f(\widetilde{x}_1 + \Delta_1) = f(1.6) = 0.76$, so we get

$$\Delta = |f(\widetilde{x}_1 + \Delta_1) - f(\widetilde{x}_1)| = |0.24 - 0.24| = 0,$$

and the range is

$$[\widetilde{y} - \Delta, \widetilde{y} + \Delta] = [0.24 - 0, 0.24 + 0] = [0.24, 0.24].$$

Problem 8–9. Linearize the expression from Problem 5 around the intervals' midpoints. Use the linearized expression to find the approximate value of the range of the original function, both with the actual derivative and with the result of numerical differentiation.

Solution. Here,

$$\widetilde{x}_1 = \frac{0.8 + 1.2}{2} = 1.0, \quad \Delta_1 = \frac{1.2 - 0.8}{2} = 0.2,$$

$$\widetilde{x}_2 = \frac{(-1.2) + (-0.8)}{2} = -1, \quad \Delta_2 = \frac{(-0.8) - (-1.2)}{2} = \frac{-0.8 + 1.2}{2} = 0.2,$$

So,

$$\widetilde{y} = f(\widetilde{x}_1, \widetilde{x}_2) = f(1, -1) = (1 - 1) \cdot (2 - (-1)) = 0 \cdot 3 = 0.$$

If we use actual derivatives, we get

$$c_1 = \frac{\partial f}{\partial x_1}(1.0, -1.0) = 2 - (-1) = 3$$

and

$$c_2 = \frac{\partial f}{\partial x_2}(1, -1) = 1 - 1 = 0.$$

Thus,

$$\Delta = |c_1| \cdot \Delta_1 + |c_2| \cdot \Delta_2 = |3| \cdot 0.2 + |0| \cdot 0.2 = 0.6 + 0 = 0.6.$$

So, the range is

$$[\widetilde{y} - \Delta, \widetilde{y} + \Delta] = [0 - 0.6, 0 + 0.6] = [-0.6, 0.6].$$

If we use numerical differentiation, then we get

$$f(\widetilde{x}_1 + \Delta_1, \widetilde{x}_2) = f(1.2, -1) = (1 - 1.2) \cdot (2 - (-1)) = (-0.2) \cdot 3 = -0.6$$

$$f(\widetilde{x}_1, \widetilde{x}_2 + \Delta_2) = f(1, -0.8) = (1 - 1) \cdot (2 - (-0.8)) = 0 \cdot 2.8 = 0.$$

Thus,

$$\Delta = |-0.6| + |0| = 0.6 + 0 = 0.6.$$

So,

$$[\widetilde{y} - \Delta, \widetilde{y} + \Delta] = [0 - 0.6, 0 + 0.6] = [-0.6, -0.6].$$

Problem 10. Use bisection to find $\sqrt{5}$ – i.e., the solution to the equation $x^2 - 5 = 0$ – with accuracy $\varepsilon = 0.25$. Start with the interval $[0, 4]$.

Solution.

- We start with $a = 0$ and $b = 4$, since
$$f(0) = 0^2 - 5 = -5 < 0$$
and
$$f(4) = 4^2 - 3 = 16 - 3 = 31 > 0.$$
So, the initial interval $[\underline{x}, \overline{x}]$ is $[0, 4]$.

- The width $4 - 0 = 2$ of the interval $[0, 2]$ is larger than $2\varepsilon = 0.5$, so we find the value $f(m)$ for the midpoint $m = 2$. For this midpoint,
$$f(2) = 2^2 - 5 = 4 - 5 = -1 < 0,$$
so we replace $\underline{x}$ with 2, and conclude that x is in the interval $[2, 4]$.

- The width $4 - 2 = 2$ of the interval $[2, 4]$ is larger than $2\varepsilon = 0.5$, so we find the value $f(m)$ for the midpoint $m = 3$. For this midpoint,
$$f(3) = 3^2 - 5 = 9 - 5 = 4 > 0,$$
so we replace $\overline{x}$ with 3, and conclude that x is in the interval $[2, 3]$.

- The width $3 - 2 = 1$ of the interval $[2, 3]$ is larger than $2\varepsilon = 0.5$, so we find the value $f(m)$ for the midpoint $m = 2.5$. For this midpoint,
$$f(2.5) = 2.5^2 - 5 = 6.25 - 5 = 1.25 > 0,$$
so we replace $\overline{x}$ with 2.5, and conclude that x is in the interval $[2, 2.5]$.

The width $2.5 - 2 = 0.5$ of the interval $[2, 2.5]$ is smaller than or equal to $2\varepsilon = 0.5$, so we return the midpoint of this interval as the desired answer:
$$x_0 = \frac{2 + 2.5}{2} = 2.25.$$