Solutions to Final Exam for Interval Class
Fall 2023

Problem 1. Why do we need data processing? Why do we need interval computations? What is the main problem of interval computations?

Solution.

• *Why do we need data processing?*

We want to predict the future state of the world, we want to know how to charge 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

$$x_1, \ldots, x_n.$$

• *Why do we need interval computations?*

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.*

- *What is the main problem of interval computations?*

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 2–4. Use several methods to estimate the range of the function $x^2 + 2x + 1$ on the interval $[-2, 0]$.

- Use calculus.
- Why cannot we use calculus to find the range for any number of inputs?
- Use linearization method based on the actual derivative.
- Use linearization method based on numerical differentiation.
- Use straightforward interval computations.
- Use full interval computations: monotonicity checking, centered form, and bisection (one bisection step is sufficient).

Solution.

• *Use calculus.*

To find this range, we can use calculus, i.e., find the values of the function $f(x)$ at the endpoints of the interval $[-2, 0]$ and at all the points where its derivative is equal to 0. The derivative is equal to $f'(x) = 2x + 2$, so it is equal to 0 when $x = -1$.

For the endpoints, we get

$$f(0) = 0^2 + 2 \cdot 0 + 1 = 1$$

and

$$f(-2) = (-2)^2 + (-2) \cdot 2 + 1 = 4 - 4 + 1 = 1.$$

For $x = -1$, we get

$$f(1) = (-1)^2 + 2 \cdot (-1) + 1 = 1 - 2 + 1 = 0.$$

The smallest of the resulting three values (1, 1, and 0) is 0, the largest is 1. So, the range is equal to $[0, 1]$.

• *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.

• *Use linearization method based on the actual derivative.*

The midpoint of the interval is $\widetilde{x} = \dfrac{-2+0}{2} = -1$, the half-width of the interval is $\Delta_1 = \dfrac{0-(-2)}{2} = 1$. The result $\widetilde{y} = f(\widetilde{x})$ of data processing is equal to

$$\widetilde{y} = f(-1) = (-1)^2 + 2 \cdot (-1) + 1 = 1 - 2 + 1 = 0.$$

The derivative is $f'(x) = 2x+2$, so the value of the derivative at the midpoint is $c_1 = f'(\widetilde{x}) = 2 \cdot (-1) + 2 = 0$. So, $\Delta = |c_1| \cdot \Delta_1 = 0$, and the estimated range is

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

• *Use linearization method based on numerical differentiation.*

We have $\Delta = |f(\widetilde{x} + \Delta_1) - f(\widetilde{x})|$. We already know that $f(\widetilde{x}) = 0$. We have

$$f(\widetilde{x} + \Delta_1) = f(-1 + 1) = f(0) = 1,$$

thus, $\Delta = |1 - 0| = 1$, and the estimated range is $[\widetilde{y} - \Delta, \widetilde{y} + \Delta] = [-1, 1]$.

• *Use straightforward interval computations.*

We have

$$[-2, 0] \cdot [-2, 0] + 2 \cdot [-2, 0] + 1 = [0, 4] + [-4, 0] + 1 = [-4, 4] + 1 = [-3, 5].$$

• *Use full interval computations: monotonicity checking, centered form, and bisection (one bisection step is sufficient).*

First, we check monotonicity. The derivative is $f'(x) = 2x + 2$, its range is

$$f'([\underline{x}, \overline{x}] = 2 \cdot [-2, 0] + 2 = [-4, 0] + 2 = [-2, 2].$$

This range has both positive and negative values, so the function is not monotonic.

For this function, the centered form leads to

$$f(\widetilde{x}) + f'([\underline{x}, \overline{x}]) \cdot [-\Delta, \Delta] = 0 + [-2, 2] \cdot [-1, 1] = [-2, 2].$$

If we bisect, we get two intervals $[-2, -1]$ and $[-1, 0]$.

- For the interval $[-1, 0]$, the range of the derivative is $f'([-1, 0]) = 2 \cdot [-1, 0] + 2 = [-2, 0] + 2 = [0, 2]$. All these values are non-negative, so the function is increasing. Thus, its range is equal to $[f(-1), f(0)] = [0, 1]$.
- For the interval $[-2, -1]$, the range of the derivative is $f'([-2, -1]) = 2 \cdot [-2, -1] + 2 = [-4, -2] + 2 = [-2, 0]$. All these values are non-postive, so the function is decreasing. Thus, its range is equal to $[f(-1), f(-2)] = [0, 1]$.

The range of the function $f(x)$ over the whole interval $[-2, 0]$ is equal to the union of its ranges over $[-2, -1]$ and $[-1, 0]$, i.e.,

$$[0, 1] \cup [0, 1] = [0, 1].$$

Problem 5. Use the interval-based optimization algorithm to locate the maximum of the function $x^2 + 2x + 1$ on the interval $[-2, 0]$. Divide this interval into two, then divide the remaining intervals into two again, etc. At each iteration, dismiss subintervals where maximum cannot be attained. Stop when you get intervals of width 0.5.

Solution. Let us first divide the original interval $[-2, 0]$ into two subintervals: $[-2, -1]$ and $[-1, 0]$.

Interval $[-1, 0]$**.** On the second interval $[-1, 0]$, the range of the derivative $2x + 2$ is equal to

$$2 \cdot [-1, 0] + 2 = [-2, 0] + 2 = [0, 2].$$

All these values are non-negative, and all but one are positive, so the function is strictly increasing on this subinterval. Its largest value is attained when x is the largest $x = 0$ and is equal to

$$0^2 + 2 \cdot 0 + 1 = 1.$$

Interval $[-2, -1]$**.** For the first subinterval $[-2, -1]$, the range of the derivative is equal to

$$2 \cdot [-2, -1] + 2 = [-4, -2] + 2 = [-2, 0].$$

All these values are non-positive, and all but one values are negative, so the function is strictly decreasing on this subinterval. Its largest value is attained when x is the smallest $x = -2$ and is equal to

$$(-2)^2 + 2 \cdot (-2) + 1 = 4 - 4 + 1 = 1.$$

The two values are equal, so the maximum is attained at two points: $x = 0$ and $x = -2$.

Problem 6. Use constraints methods to solve the systems $x_1 + 2x_2 = 1$ and $x_1 - 2x_2 = 0$ for $x_1, x_2 \in [0,1]$. Use one bisection step.

Solution. Based on the equations, we form the following four rules:

1. $x_1 \mapsto x_2 = 0.5 - 0.5 \cdot x_1$;
2. $x_2 \mapsto x_1 = 1 - 2x_2$;
3. $x_1 \mapsto x_2 = 0.5 \cdot x_1$;
4. $x_2 \mapsto x_1 = 2x_2$.

We start with $x_1 \in \mathbf{x}_1 = [0,1]$ and $x_2 \in \mathbf{x}_2 = [0,1]$. In general, we apply all the rules again and again. If we get stuck, we bisect one of the intervals.

- Rule 1 leads to $\mathbf{x}_2 = (0.5 - [0,0.5]) \cap [0,1] = [0,0.5] \cap [0,1] = [0,0.5]$.
- Rule 2 leads to $\mathbf{x}_1 = (1 - 2 \cdot [0,1]) \cap [0,1] = [-1,1] \cap [0,1] = [0,1]$.
- Rule 3 leads to $\mathbf{x}_2 = [0,0.5] \cap [0,0.5] = [0,0.5]$.
- Rule 4 leads to $\mathbf{x}_1 = [0,1] \cap [0,1] = [0,1]$.
- Rule 1 leads to $\mathbf{x}_2 = (0.5 - [0,0.5]) \cap [0,1] = [0,0.5] \cap [0,1] = [0,0.5]$.

We had a full cycle, and the intervals did not change. So, we bisect one of the intervals – e.g., we bisect the interval $\mathbf{x}_1 = [0,1]$ into intervals $[0,0.5]$ and $[0.5,1]$.

Let us first consider the first of the new intervals, i.e., the case when $\mathbf{x}_1 = [0,0.5]$ and $\mathbf{x}_2 = [0,0.5]$.

- Rule 1 leads to $\mathbf{x}_2 = (0.5 - [0,0.25]) \cap [0,0.5] = [0.25,0.5] \cap [0,0.5] = [0.25,0.5]$.
- Rule 2 leads to $\mathbf{x}_1 = (1 - 2 \cdot [0.5,1]) \cap [0,0.5] = [0,1] \cap [0,0.5] = [0,0.5]$.
- Rule 3 leads to $\mathbf{x}_2 = (0.5 \cdot [0,0.5]) \cap [0.25,0.5] = [0,0.25] \cap [0.25,0.5] = [0.25,0.25]$.
- Rule 4 leads to $\mathbf{x}_1 = (2 \cdot [0.25,0.25]) \cap [0,0.5] = [0.5,0.5] \cap [0,0.5] = [0.5,0.5]$.

So, the solution is $x_1 = 0.5$ and $x_2 = 0.25$.

Let us now consider the second of the new intervals, i.e., the case when $\mathbf{x}_1 = [0.5,1]$ and $\mathbf{x}_2 = [0,0.5]$.

- Rule 1 leads to $\mathbf{x}_2 = (0.5 - [0.25,0.5]) \cap [0,0.5] = [0,0.25] \cap [0,0.5] = [0,0.25]$.
- Rule 2 leads to $\mathbf{x}_1 = (1 - 2 \cdot [0,0.25]) \cap [0.5,1] = [0.5,1] \cap [0.5,1] = [0.5,1]$.
- Rule 3 leads to $\mathbf{x}_2 = (0.5 \cdot [0.5,1]) \cap [0,0.25] = [0.25,0.5] \cap [0,0.25] = [0.25,0.25]$.
- Rule 4 leads to $\mathbf{x}_1 = (2 \cdot [0.25,0.25]) \cap [0.5,1] = [0.5,0.5] \cap [0.5,1] = [0.5,0.5]$.

So, we get the same solution: $x_1 = 0.5$ and $x_2 = 0.25$.

Problem 7. If we have $\mu(1) = 0.6$ and $\mu(3) = 0$, what value shall we assign to $\mu(2)$? Use linear interpolation.

Solution. Linear interpolation means

$$f(x) = y_1 + \frac{y_2 - y_1}{x_2 - x_1} \cdot (x - x_1).$$

In our case, $x_1 = 1$, $y_1 = 0.6$, $x_2 = 3$, $y_2 = 0$, and $x = 2$. So, we get

$$\mu(2) = 0.6 + \frac{0 - 0.6}{3 - 1} \cdot (2 - 1) = 0.6 - \frac{0.6}{2} \cdot 1 = 0.6 - 0.3 = 0.3.$$

Problem 8. If for $\alpha = 0.7$, the α-cut of x is $[-2, 0]$, and $y = x^2 + 2x + 1$, what is the corresponding α-cut of y? Explain your answer.

Solution. The α-cut of y is the range $f([-2, 0])$ of the function $f(x) = x^2 + 2x + 1$ on the interval $[-2, 0]$. To find this range, we can use calculus, i.e., find the values of the function $f(x)$ at the endpoints of the interval $[-2, 0]$ and at all the points where its derivative is equal to 0. The derivative is equal to $f'(x) = 2x + 2$, so it is equal to 0 when $x = -1$.

For the endpoints, we get $f(0) = 0^2 + 2 \cdot 0 + 1 = 1$ and $f(2) = (-2)^2 + 2 \cdot (-2) + 1 = 4 - 4 + 1$. For $x = -1$, we get $f(-1) = (-1)^2 + 2 \cdot (-1) + 1 = 1 - 2 + 1 = 0$. The smallest of the resulting three values 1, 1, and 0 is 0, the largest is 1. So, the desired range is equal to $[0, 1]$.

Thus, the desired α-cut is also equal to $[0, 1]$.