Why Linear Interpolation?

Andrzej Pownuk and Vladik Kreinovich

Computational Science Program
University of Texas at El Paso
500 W. University
El Paso, Texas 79968, USA
ampownuk@utep.edu, vladik@utep.edu

1. Need for Interpolation

- In many practical situations:
  - we know that the value of a quantity y is uniquely determined by the value of some other quantity x,
  - but we do not know the exact form of the corresponding dependence $y = f(x)$.
- To find this dependence, we measure the values of x and y in different situations.
- As a result, we get the values $y_i = f(x_i)$ of the unknown function $f(x)$ for several values $x_1, \ldots, x_n$.
- Based on this information, we would like to predict the value $f(x)$ for all other values x.
- When x is between the smallest and the largest of the values x_i, this prediction is known as the *interpolation*.

2. Why Linear Interpolation?

- Let's consider the case $n = 2$. Let's assume that $f(x)$ is linear on $[x_1, x_2]$; then

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

- This formula is known as *linear interpolation.*
- The usual motivation for linear interpolation is simplicity: linear functions are the easiest to compute.
- An interesting empirical fact is that in many practical situations, linear interpolation works reasonably well.
- We know that in computational science, often very complex computations are needed.
- So we cannot claim that nature prefers simplicity.
- There should be another reason for the empirical fact that linear interpolation often works well.

3. Reasonable Properties of Interpolation

- We want to be able,
  - given values y_1 and y_2 of the unknown function at points x_1 and x_2, and a point $x \in (x_1, x_2)$,
  - to provide an estimate for $f(x)$.
- Let us denote this estimate by $I(x_1, y_1, x_2, y_2, x)$; what are the reasonable properties of this function?
- If $y_i = f(x_i) \le y$ for both i, it is reasonable to expect that $f(x) \le y$.
- In particular, for $y = \max(y_1, y_2)$, we conclude that

$$I(x_1, y_1, x_2, y_2, x) \le \max(y_1, y_2).$$

- Similarly, if $y \le y_i$ for both i, it is reasonable to expect that $y \le f(x)$.
- In particular, for $y = \min(y_1, y_2)$, we conclude that

$$\min(y_1, y_2) \le I(x_1, y_1, x_2, y_2, x).$$

4. x-Scale-Invariance

- The numerical value of a physical quantity depends:
  - on the choice of the measuring unit and
  - on the starting point.
- If we change the starting point to the one which is b units smaller, then b is added to all the values.
- If we replace a measuring unit by a $a > 0$ times smaller one, then all the values are multiplied by a.
- If we perform both changes, then each original value x is replaced by the new value $x' = a \cdot x + b$.
- For example, if we know the temperature x in C, then the temperature x' in F is $x' = 1.8 \cdot x + 32$.
- The interpolation procedure should not change if we simply re-scale:

  $$I(a \cdot x_1 + b, y_1, a \cdot x_2 + b, y_2, a \cdot x + b) = I(x_1, y_1, x_2, y_2, x).$$

5. y-Scale-Invariance

- Similarly, we can consider different units for y.
- The interpolation result should not change if we simply change the starting point and the measuring unit; so:
  - if we replace y_1 with $a \cdot y_1 + b$ and y_2 with $a \cdot y_2 + b$,
  - then the result of interpolation should be obtained by a similar transformation from the previous one:

$$I(x_1, a \cdot y_1 + b, x_2, a \cdot y_2 + b, x) = a \cdot I(x_1, y_1, x_2, y_2, x) + b.$$

6. Consistency

- When $x_1 \le x_1' \le x \le x_2' \le x_2$, the value $f(x)$ can be estimated in two different ways.
- We can interpolate directly from the values $y_1 = f(x_1)$ and $y_2 = f(x_2)$, getting $I(x_1, y_1, x_2, y_2, x)$.
- Or we can:
  - first estimate the values $f(x_1') = I(x_1, y_1, x_2, y_2, x_1')$ and $f(x_2') = I(x_1, y_1, x_2, y_2, x_2')$, and
  - then use these two estimates to estimate $f(x)$ as

$$I(x_1, f(x_1'), x_2, f(x_2'), x) =$$

$$I(x_1', I(x_1, y_1, x_2, y_2, x_1'), x_2', I(x_1, y_1, x_2, y_2, x_2'), x).$$

- It is reasonable to require that these two ways lead to the same estimate for $f(x)$: $I(x_1, y_1, x_2, y_2, x) =$

$$I(x_1', I(x_1, y_1, x_2, y_2, x_1'), x_2', I(x_1, y_1, x_2, y_2, x_2'), x).$$

7. Continuity

- Most physical dependencies are continuous.
- Thus, when the two value x and x' are close, we expect the estimates for $f(x)$ and $f(x')$ to be also close.
- Thus, it is reasonable to require that:
  - the interpolation function $I(x_1, y_1, x_2, y_2, x)$ is continuous in x, and
  - that for both $i = 1, 2$, $I(x_1, y_1, x_2, y_2, x)$ converges to $f(x_i)$ when $x \to x_i$.

8. Resulting Definition

A function $I(x_1, y_1, x_2, y_2, x)$ defined for $x_1 < x < x_2$ is called an *interpolation function* if:

- $\min(y_1, y_2) \le I(x_1, y_1, x_2, y_2, x) \le \max(y_1, y_2)$;
- $I(a \cdot x_1 + b, y_1, a \cdot x_2 + b, y_2, a \cdot x + b) = I(x_1, y_1, x_2, y_2, x)$ for all x_i, y_i, x, $a > 0$, and b (x-scale-invariance);
- $I(x_1, a \cdot y_1 + b, x_2, a \cdot y_2 + b, x) = a \cdot I(x_1, y_1, x_2, y_2, x) + b$ for all x_i, y_i, x, $a > 0$, and b (y-scale invariance);
- consistency: $I(x_1, y_1, x_2, y_2, x) =$

$$I(x'_1, I(x_1, y_1, x_2, y_2, x'_1), x'_2, I(x_1, y_1, x_2, y_2, x'_2), x);$$

- continuity:
  - the expression $I(x_1, y_1, x_2, y_2, x)$ is a continuous function of x,
  - $I(x_1, y_1, x_2, y_2, x) \to y_1$ when $x \to x_1$ and $I(x_1, y_1, x_2, y_2, x) \to y_2$ when $x \to x_2$.

9. Main Result

- **Result:** *The only interpolation function satisfying all the properties is the linear interpolation*

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

- Thus, we have indeed explained that linear interpolation follows from the fundamental principles.
- This may explain its practical efficiency.

10. Proof

- When $y_1 = y_2$, the conservativeness property implies that $I(x_1, y_1, x_2, y_1, x) = y_1$.
- Thus, to complete the proof, it is sufficient to consider two remaining cases: when $y_1 < y_2$ and when $y_2 < y_1$.
- We will consider the case when $y_1 < y_2$.
- The case when $y_2 < y_1$ is considered similarly.
- So, in the following text, without losing generality, we assume that $y_1 < y_2$.

11. Using y-Scale-Invariance

- When $y_1 < y_2$, then $y_1 = a \cdot 0 + b$ and $y_2 = a \cdot 1 + b$ for $a = y_2 - y_1$ and y_1.
- Thus, the y-scale-invariance implies that

$$I(x_1, y_1, x_2, y_2, x) = (y_2 - y_1) \cdot I(x_1, 0, x_2, 1, x) + y_1.$$

- If we denote $J(x_1, x_2, x) \stackrel{\text{def}}{=} I(x_1, 0, x_2, 1, x)$, then we get

$$I(x_1, y_1, x_2, y_2, x) = (y_2 - y_1) \cdot J(x_1, x_2, x) + y_1 =$$

$$J(x_1, x_2, x) \cdot y_2 + (1 - J(x_1, x_2, x)) \cdot y_1.$$

12. Using x-Scale-Invariance

- Since $x_1 < x_2$, we have $x_1 = a \cdot 0 + b$ and $x_2 = a \cdot 1 + b$, for $a = x_2 - x_1$ and $b = x_1$.
- Here, $x = a \cdot r + b$, where $r = \dfrac{x - b}{a} = \dfrac{x - x_1}{x_2 - x_1}$.
- Thus, the x-scale invariance implies that $J(x_1, x_2, x) = w\left(\dfrac{x - x_1}{x_2 - x_1}\right)$, where $w(r) \stackrel{\text{def}}{=} J(0, 1, r)$.
- Thus, the above expression for $I(x_1, y_1, x_2, y_2, x)$ in terms of $J(x_1, x_2, x)$ takes the following simplified form:

$$w\left(\frac{x - x_1}{x_2 - x_1}\right) \cdot y_2 + \left(1 - w\left(\frac{x - x_1}{x_2 - x_1}\right) \cdot y_2\right) \cdot y_1.$$

- To complete our proof, we need to show that $w(r) = r$ for all $r \in (0, 1)$.

13. Using Consistency

- Let us take $x_1 = y_1 = 0$ and $x_2 = y_2 = 1$, then $I(0,0,1,1,x) = w(x) \cdot 1 + (1 - w(x)) \cdot 0 = w(x)$.
- For $x = 0.25 = \dfrac{0 + 0.5}{2}$, the value $w(0.25)$ can be obtained by interpolating $w(0) = 0$ and $\alpha \stackrel{\text{def}}{=} w(0.5)$:
$$w(0.25) = \alpha \cdot w(0.5) + (1 - \alpha) \cdot w(0) = \alpha^2.$$
- For $x = 0.75 = \dfrac{0.5 + 1}{2}$, we similarly get:
$$w(0.75) = \alpha \cdot w(1) + (1 - \alpha) \cdot w(0.5) = \alpha \cdot 1 + (1 - \alpha) \cdot \alpha = 2\alpha - \alpha^2.$$
- $w(0.5)$ can be interpolated from $w(0.25)$ and $w(0.75)$:
$$w(0.5) = \alpha \cdot w(0.75) + (1 - \alpha) \cdot w(0.25) =$$
$$\alpha \cdot (2\alpha - \alpha^2) + (1 - \alpha) \cdot \alpha^2 = 3\alpha^2 - 2\alpha^3.$$
- By consistency, this estimate should be equal to our original estimate $w(0.5) = \alpha$: $3\alpha^2 - 2\alpha^3 = \alpha$.

14. What Is α

- Here, $\alpha = w(0.5) = 0$, $\alpha = 1$, or $\alpha = 0.5$.
- If $\alpha = 0$, then, $w(0.75) = \alpha \cdot w(1) + (1 - \alpha) \cdot w(0.5) = 0$.
- By induction, we can show that $\forall n\, (w(1 - 2^{-n}) = 0)$ for each n.
- Here, $1 - 2^{-n} \to 1$, but $w(1 - 2^{-n}) \to 0$, which contradicts to continuity $w(1 - 2^{-n}) \to w(1) = 1$.
- Thus, $\alpha = 0$ is impossible.
- When $\alpha = w(0.5) = 1$, then

$$w(0.25) = \alpha \cdot w(0.5) + (1 - \alpha) \cdot w(0) = 1.$$

- By induction, $w(2^{-n}) = 1$ for each n.
- In this case, $2^{-n} \to 0$, but $w(2^{-n}) \to 1$, which contradicts to continuity $w(2^{-n}) \to w(0) = 0$.
- Thus, $\alpha = 0.5$.

15. Proof: Final Part

- For $\alpha = 0.5$: $w(0) = 0$, $w(0.5) = 0.5$, $w(1) = 1$.
- Let us prove, by induction over q, that for every binary-rational number $r = \dfrac{p}{2^q} \in [0, 1]$, we have $w(r) = r$.
- Indeed, the base case $q = 1$ is proven.
- Let us assume that we have proven it for $q = 1$.
- If p is even $p = 2k$, then $\dfrac{2k}{2^q} = \dfrac{k}{2^{q-1}}$, so the desired equality comes from the induction assumption.
- If $p = 2k + 1$, then $r = \dfrac{p}{2^q} = \dfrac{2k+1}{2^q} =$

$$0.5 \cdot \frac{2k}{2^q} + 0.5 \cdot \frac{2 \cdot (k+1)}{2^q} = 0.5 \cdot \frac{k}{2^{q-1}} + 0.5 \cdot \frac{k+1}{2^{q-1}}.$$

- So $w(r) = 0.5 \cdot w\left(\dfrac{k}{2^{q-1}}\right) + 0.5 \cdot w\left(\dfrac{k+1}{2^{q-1}}\right)$.

16. Proof: Final Part (cont-d)

- By induction assumption, we have

$$w\left(\frac{k}{2^{q-1}}\right)=\frac{k}{2^{q-1}} \text{ and } w\left(\frac{k+1}{2^{q-1}}\right)=\frac{k+1}{2^{q-1}}.$$

- Thus, $w(r)=\alpha\cdot\dfrac{k}{2^{q-1}}+0.5\cdot\dfrac{k+1}{2^{q-1}}=\dfrac{2k+1}{2^q}=r.$
- The equality $w(r)=r$ is hence true for all binary-rational numbers.
- Any real number x from the interval $[0,1]$ is a limit of such numbers – truncates of its binary expansion.
- Thus, by continuity, we have $w(x)=x$ for all x.
- Substituting $w(x)=x$ into the above formula for $I(x_1,y_1,x_2,y_2,x)$ leads to linear interpolation. Q.E.D.

17. Acknowledgments

This work was supported in part:

- by the National Science Foundation grants:
  - HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and
  - DUE-0926721, and
- by an award from Prudential Foundation.