Introduction to successive iteration and open the door to numerical methods with the square root example. This worksheet includes several animation clips on successive iteration. Instructor: Nam Sun Wang

Successive iteration is extremely common in numerical computations whether we are trying to find a solution to a (set of) linear or nonlinear algebraic equations, matrix inverse, or ordinary and partial differential equations. You also see it in probability transition matrix, chaos, and the generation of fractal patterns. The iterative scheme takes the general form of

x=g(x)

The above form allows us to start with an initial value of x, plug it into function g, and the function g will provide us with the next value of x. We insert the new value of x back into f(x) again, which will crank out yet another x for us. We can repeat the process for as many times as we wish. This process is called **successive iteration** or **successive approximation** (in cases where we resort to iteration to compensate for approximation). Note that the above successive iteration scheme contains a purely x term on the LHS. Let us illustrate the successive methods with the old-fashioned square root problem where the objective is to find a number x such that

 $x^2 = x \cdot x = a$ (1)

Note that algebra tells us that there are two roots. In the following discussion, let us take as an example:

a := 10

Of course, we symbolically denote such a number as:

$$\sqrt{a}$$
 or $\sqrt{10}$ (And the other number as $-\sqrt{a}$ or $-\sqrt{10}$

But what is its value? Here, we want to find the square root without resorting to the square root key on our calculator or without calling the built-in square root operator/function in a numerical package. We recall that 3*3=9 and 4*4=16, thus the answer should lie between 3 and 4. Let's start with:

)

$$x_0 = 3$$

We have to add a little bit of correction to x_0 . Let this correction be ε_0 . The updated root after making the correction is:

 $x_1 = x_0 + \varepsilon_0 = \sqrt{a}$

To find ε_0 , we square both sides of the last equation. Remember, we are after a number whose square is *a*.

$$(x_1)^2 = (x_0 + \varepsilon_0)^2 = a$$

Expanding (mark the last equation and choose |Symbolic|Expand Expression| from Mathcad menu) yields:

$$(\mathbf{x}_1)^2 = (\mathbf{x}_0)^2 + 2 \cdot \mathbf{x}_0 \cdot \mathbf{\varepsilon}_0 + (\mathbf{\varepsilon}_0)^2 = \mathbf{a} \longrightarrow (\mathbf{\varepsilon}_0)^2 + 2 \cdot \mathbf{x}_0 \cdot \mathbf{\varepsilon}_0 + (\mathbf{x}_0)^2 - \mathbf{a} = 0$$
 (2)

Substituting the values of $x_0=3$ and a=10, we get:

$$\left(\varepsilon_{0}\right)^{2} + 2 \cdot 3 \cdot \varepsilon_{0} + 3^{2} - 10 = 0 \qquad \longrightarrow \qquad \left(\varepsilon_{0}\right)^{2} + 6 \cdot \varepsilon_{0} - 1 = 0$$

Recall the well known quadratic formula to solve for ε_0 :

$$\epsilon_0 = \frac{-6 - \sqrt{6^2 + 4}}{2} = -3 - \sqrt{10}$$
 and $\epsilon_0 = \frac{-6 + \sqrt{6^2 + 4}}{2} = -3 + \sqrt{10}$

Ocops. We end up with an expression that contains the square root of 10, which we originally set out to solve in the first place. The problem arises because we are trying to solve the quadratic equation (2) in ε *exactly* with the quadratic formula. Since this leads us going in circles, we propose to **make an approximation** -- the key word in numerical method! Let us drop the ε_0^2 term from Equation (2).

$$2 \cdot x_0 \cdot \varepsilon_0 + (x_0)^2 - a=0$$
 (The notation is a bit sloppy here, because I cannot find the approximation sign in Mathcad.)

With this approximation, we have a much more manageable equation that does not require the quadratic formula to find ε_0 :

$$\varepsilon_0 = \frac{a - (x_0)^2}{2 \cdot x_0}$$

Let us actually plug some numbers into this formula.

$$\varepsilon_0 := \frac{\mathbf{a} - \left(\mathbf{x}_0\right)^2}{2 \cdot \mathbf{x}_0} \qquad \varepsilon_0 = 0.167$$

The next value of x is:

$$x_1 := x_0 + \varepsilon_0$$
 $x_1 = 3.167$ Check: $(x_1)^2 = 10.028$ \leftarrow not quite 10.

Since we have made an approximation by dropping the ε_0^2 term, the correction we have just made is not an exact one, as we can see from the check Nevertheless, we are now a step closer to the actual root. **The price to pay for making an approximation is iteration.** Now, x_1 is a better estimate of sqrt(10), let us make another correction ε_1 on top of x_1 . We hope the resulting estimate x_2 will be an even better one.

 $x_2 = x_1 + \varepsilon_1$

To find ε_1 , we follow the same set of steps as before: square both sides of the last equation, equate it to *a*, and expand the square term.

$$(x_2)^2 = (x_1 + \varepsilon_1)^2 = a$$

$$(x_1)^2 + 2 \cdot x_1 \cdot \varepsilon_1 + (\varepsilon_1)^2 = a$$

$$(\varepsilon_1)^2 + 2 \cdot x_1 \cdot \varepsilon_1 + (x_1)^2 - a = 0$$

As before, we make an approximating by dropping the ε_1^2 term.

$$2 \cdot \mathbf{x}_{1} \cdot \mathbf{\varepsilon}_{1} + (\mathbf{x}_{1})^{2} - \mathbf{a} = 0$$
$$\mathbf{\varepsilon}_{1} = \frac{\mathbf{a} - (\mathbf{x}_{1})^{2}}{2 \cdot \mathbf{x}_{1}}$$

Plug in some numbers.

$$\varepsilon_1 := \frac{\mathbf{a} - (\mathbf{x}_1)^2}{2 \cdot \mathbf{x}_1} \qquad \varepsilon_1 = -0.004$$

The next value of x is:

$$x_2 := x_1 + \varepsilon_1$$
 $x_2 = 3.16228$ Check: $(x_2)^2 = 10.00002 \leftarrow \text{almost 10}$.

If we are satisfied with the results, we stop here. Otherwise, repeat until we are satisfied with the accuracy. We shall do it just one more time, this time without derivation.

$$\varepsilon_2 := \frac{\mathbf{a} - (\mathbf{x}_2)^2}{2 \cdot \mathbf{x}_2} \qquad \varepsilon_2 = -3.042 \cdot 10^{-6}$$

The next value of x is:

Mathcad's build-in function gives: $\sqrt{10} = 3.162277660168$ which agrees with our value to more than 10 digits.

The general iteration scheme for square root of a is:

For N := 5 i := 0.. N and provide
$$x_0$$
,
 $x_{i+1} := x_i + \frac{a - (x_i)^2}{2 \cdot x_i}$

Although we have started with a reasonably close initial guess of 3, this does not have to be so. We can see that any nonzero initial guess will eventually lead to a root with this scheme. (We may need more number of iterations though.) Zero is not a good initial guess because of the problem with division by 0.

For N = 15 i = 0... N $x_0 = 1000 \leftarrow A$ lousy initial guess. $x_{i+1} := x_i + \frac{a - (x_i)^2}{2 \cdot x_i}$ $x_3 = 125.0262489500585 \leftarrow \text{Not quite there yet.}$ $x_{last(x)} = 3.16227766016838 \leftarrow Good!$

Furthermore, a different initial guess may lead to a different root when multiple roots exist. Of course, we know that x*x=a has two solutions, a positive one and a negative one, although sqrt(a) refers to the positive root and -sqrt(a) refers to the negative one.

 $x_0 := -1000 \quad \leftarrow$ Another lousy initial guess from the negative side.

$$\begin{array}{ll} x_{i+1} \coloneqq x_i + \frac{a - \left(x_i\right)^2}{2 \cdot x_i} & x_3 = -125.0262489500585 & \leftarrow \text{Not quite there yet.} \\ & x_{\text{last}(x)} = -3.16227766016838 \leftarrow \text{Good, but a negative one.} \end{array}$$

The same formula is also applicable to different values of *a*.

a := 100

$$x_0 := 1$$

 $x_{i+1} := x_i + \frac{a - (x_i)^2}{2 \cdot x_i}$ $x_{last(x)} = 10.00000000000 \leftrightarrow \text{Good!}$

In terms of the general iteration form of x=g(x) mentioned at the beginning of this worksheet, the iteration fomula is:

$$g(x) := x + \frac{a - x^2}{2 \cdot x}$$
(3)
$$x_{i+1} := g(x_i)$$

In Mathcad version 6 (not valid in version 5), we can iterate elegantly by defining a function recursively. The following function says if x=g(x) is satisfied, return x and stop; otherwise call g(x) to update x and iterate until convergence. The intended usage is to issue the initial guess as the argument to the "iterate" function.

$$iterate(x) := (x = g(x)) \cdot x + (x \neq g(x)) \cdot iterate(g(x))$$

iterate(3) = ... display the value returned with an initial guess of 3

Another way of saying the same thing:

iterate(x) := if(x=g(x), x, iterate(g(x))) $iterate(3) = \dots display the value returned with an initial guess of 3$

5

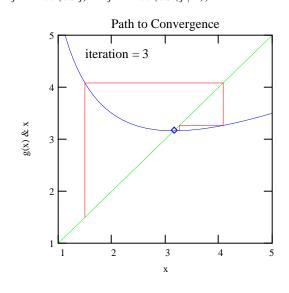
iterate.mcd

Visualization of Successive Iteration x=g(x). The solution is where the iteration function g(x) intersects with the diagonal line, which is a straight function of x.

a := 10
$$g(x) := \frac{x^2 + a}{2 \cdot x}$$
 N := 3 i := 0.. N
x₀ := 1.5 $x_{i+1} := g(x_i)$

Generate the steps for plotting:

$$u_j = x_{floor(0.5 \cdot j)}$$
 $v_j = x_{floor(0.5 \cdot (j+1))}$ $xx = 1, 1.1 .. 5$



Animation section: toggle off the next equation and set FRAME=0..2N

FRAME = $2 \cdot N$ FRAME = 6

j = 0 .. FRAME

iteration = floor($0.5 \cdot FRAME$)

Click on the following icon to play an animation clip.



iterate1.avi

Let us now follow similar steps to find a **cubic root** of a given number *a*. The objective is to find x such that $x^*x^*x=a$. Approximate x first, then make a correction.

$$a := 10 \leftarrow A$$
 given number whose cubic root is to be estimated.

 $x_0 := 2 \leftarrow \text{Initial guess.}$

$$x_1 = x_0 + \varepsilon_0$$

We calculate the correction term ϵ_0 from (|Symbolic|Expand Expression| with Mathcad):

$$\begin{aligned} & \left(x_{1}\right)^{3} = \left(x_{0} + \varepsilon_{0}\right)^{3} = a \\ & \left(x_{0}\right)^{3} + 3 \cdot \left(x_{0}\right)^{2} \cdot \varepsilon_{0} + 3 \cdot x_{0} \cdot \left(\varepsilon_{0}\right)^{2} + \left(\varepsilon_{0}\right)^{3} = a \end{aligned}$$

Ignore the quadratic term (ϵ_0^2) and the cubic term (ϵ_0^3):

$$(x_0)^3 + 3 \cdot (x_0)^2 \cdot \varepsilon_0 = a$$
$$\varepsilon_0 = \frac{a - (x_0)^3}{3 \cdot (x_0)^2}$$

Plug in some numbers.

$$\varepsilon_0 := \frac{a - (x_0)^3}{3 \cdot (x_0)^2} \qquad \varepsilon_0 = 0.167$$

Update x:

$$x_1 := x_0 + \varepsilon_0$$
 $x_1 = 2.167$ Check: $(x_1)^3 = 10.171$

Repeat the correction a couple more times:

$$\begin{split} \epsilon_{1} &:= \frac{a - (x_{1})^{3}}{3 \cdot (x_{1})^{2}} & x_{2} := x_{1} + \epsilon_{1} & x_{2} = 2.1545 & \text{Check:} \quad (x_{2})^{3} = 10.001 \\ \epsilon_{2} &:= \frac{a - (x_{2})^{3}}{3 \cdot (x_{2})^{2}} & x_{3} := x_{2} + \epsilon_{2} & x_{3} = 2.15443469 & \text{Check:} \quad (x_{3})^{3} = 10.00000003 \\ & \text{Mathcad's built-in routine gives:} \quad a^{\frac{1}{3}} = 2.15443469 \end{split}$$

It is clear that the scheme is converging rapidly to an actual cubic root. Thus, the **general iteration** scheme for a cubic root of *a* is:

For N = 15 i = 0.. N and provide an initial guess x_0 ,

$$x_{i+1} := x_i + \frac{a - (x_i)^3}{3 \cdot (x_i)^2}$$

We know from mathematical theories that there are three roots to a cubic equation. For $x^3=a$, where *a* is a real number, there is a real root and two complex roots. To find a complex root, we must start with a complex guess:

$$a := 1 \qquad x_0 := 1 + -0.7i \qquad \leftarrow \text{ Initial guess. The "i" here denotes the imaginary part, which is entered by typing "1i", which is not to be confused with the running index "i" below. The resulting answer is very sensitive to the initial guess. In fact, we can generate an interesting fractal pattern based on whether the resulting answer is the real number "1" or a complex one.$$

 $x_{last(x)} = 1$

The general iteration formula for the n-root of a given number a is:

$a - x^n$	Example:	n :=4	a := 10	with an initial guess of $x_0 = 2$
$g(x) := x + \frac{a - x^n}{n \cdot x^{n-1}}$	yields the following answer:			
$\mathbf{x}_{i+1} := \mathbf{g}(\mathbf{x}_i)$		$x_{last(x)} =$	1.7782794	1 <u>1</u>
	Mathcad's bui	It-in routin	e gives:	$a^n = 1.77827941$

Below, we present a different way of deriving an iteration formula for finding the square root of a given number *a*. We start with what we are trying to achieve.

 $x^2 = a$

Let us add x^2 to both sides of the equation.

$$x^{2} + x^{2} = x^{2} + a$$

 $2 \cdot x^{2} = x^{2} + a$

Divide the above equation by 2x so that we obtain an equation in the form of x=g(x) suitable for iteration

$$x = \frac{x^2 + a}{2 \cdot x}$$

Thus, the iteration scheme is:

$$g(x) := \frac{x^2 + a}{2 \cdot x}$$
 (4)

Now, let us apply this iteration scheme to a = 10

Start with:
$$x_0 := 3$$

 $x_1 := g(x_0)$ $x_1 = 3.167$
 $x_2 := g(x_1)$ $x_2 = 3.162280701754$
: $x_{i+1} := g(x_i)$ $x_{last(x)} = 3.162277660168$

The algebraic equation f(x) to be solved for is: $x^2 - a=0$

$$f(\mathbf{x}) := \mathbf{x}^2 - \mathbf{a} \qquad f(\mathbf{x}) := 2 \cdot \mathbf{x}$$
$$\mathbf{x}_{i+1} := \mathbf{x}_i - \frac{f(\mathbf{x}_i)}{f(\mathbf{x}_i)}$$

Many numerical methods exist for solving nonlinear equation of the form f(x)=0. Each one of these iterate based on the same general form of x=g(x) but a different specific functional form of g(x). Thus, deriving a good numerical algorithm is to find an expression of x=g(x). Who is to say that we cannot derive a different formula? There are infinite possibilities. For example, if we add a $2x^2$ term, instead of x^2 , to both sides of the equation at the beginning of this section, the resulting iteration formula would be:

$$g(x) := \frac{2 \cdot x^2 + a}{3 \cdot x}$$

Now, let's apply this new iteration scheme to a := 10

Start with:
$$x_0 := 3$$

 $x_1 := g(x_0)$ $x_1 = 3.111$
 $x_2 := g(x_1)$ $x_2 = 3.145502645503$
:
 $x_{i+1} := g(x_i)$ $x_{last(x)} = 3.162277656689$

which is also eventually converging to the value of $\sqrt{10} = 3.162277660168$

iterate.mcd

Let us try another scheme by simply adding a term x to the both sides of equation (1). This is a favorite trick for many because it quickly converts an algebraic equation of the form 0=f(x) into a successive iteration form x=g(x), which is simply x+f(x).

$$x + x^{2} = x + a$$

 $x = x + a - x^{2}$
 $g(x) := x + a - x^{2}$ (5)

Now, let us apply this new iteration scheme to a = 10

Start with:
$$x_0 := 3$$

 $x_1 := g(x_0)$ $x_1 = 4$
 $x_2 := g(x_1)$ $x_2 = -2$
 $x_3 := g(x_2)$ $x_3 = 4$
:
 $x_{i+1} := g(x_i)$ $x_{last(x)} = -2$

Hmm... The value of x is just switching back and forth, not at all converging.

Plot for Switching Case -- Equation (5)

N := 5

$$i := 0..N$$

 $x_0 := 3$ $x_{i+1} := g(x_i)$

Generate the steps for plotting:

Animation section: toggle off the next equation and set FRAME=0..2N

FRAME = $2 \cdot N$ FRAME = 10

j := 0 .. FRAME

iteration = $floor(0.5 \cdot FRAME)$

Click on the following icon to play an animation clip.



iterate2.avi

Let us try another scheme. If we add 0.95*x to both sides of equation (1), we get:

$$\begin{array}{l} 0.95 \cdot x + x^{2} = 0.95 \cdot x + a \\ x = x + \frac{a - x^{2}}{0.95} \\ g(x) := x + \frac{a - x^{2}}{0.95} \\ (6) \\ \end{array}$$
Start with: $x_{0} := 3$
 $x_{1} := g(x_{0})$ $x_{1} = 4.053$
 $x_{2} := g(x_{1})$ $x_{2} = -2.709$
 $x_{3} := g(x_{2})$ $x_{3} = 0.090$
 $x_{4} := g(x_{3})$ $x_{4} = 10.608$ \leftarrow Oscillating away from the solution.

Hmm... The value of x is getting larger and larger and not converging into a single number. This is called **divergence**.

Plot for Diverging Case -- Equation (6)

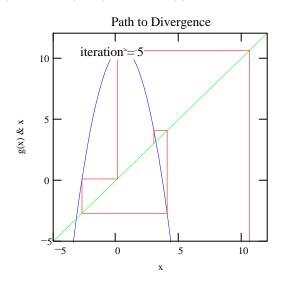
N := 5

$$x_0 := 3$$

 $x_{i+1} := g(x_i)$

Generate the steps for plotting:

$$u_j := x_{floor(0.5 \cdot j)}$$
 $v_j := x_{floor(0.5 \cdot (j+1))}$ $xx := -5, -4.9..12$



Animation section: toggle off the next equation and set FRAME=0..2N $FRAME := 2 \cdot N$ FRAME = 10

j = 0 .. FRAME

iteration = floor($0.5 \cdot FRAME$)

Click on the following icon to play an animation clip.



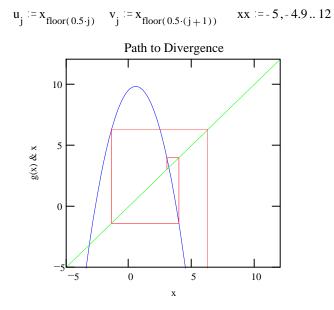
iterate3.avi

Now, if we add 1.05^*x to both sides of equation (1), we get:

$$\begin{array}{ll} g(x) \coloneqq x + \frac{a - x^2}{1.05} & (7) & \leftarrow \text{Try different values for the denominator, in place of 1.05.} \\ \begin{array}{ll} \text{Start with:} & x_0 \coloneqq 3 \\ x_1 \coloneqq g(x_0) & x_1 = 3.952 \\ x_2 \coloneqq g(x_1) & x_2 = -1.401 \\ x_3 \coloneqq g(x_2) & x_3 = 6.253 \\ x_4 \coloneqq g(x_3) & x_4 = -21.456 \\ & \leftarrow \text{Getting away from the solution.} \\ \vdots \\ x_{i+1} \coloneqq g(x_i) & x_{last(x)} = -2 \end{array}$$

11

Plot for Diverging Case -- Equation (7)



12

iterate.mcd

Another easy way of deriving an iteration expression is to reduce the given algebraic equation first into the form of f(x)=0. This is also a favorite step for many. Since f(x)=0, we can multiply, divide or do almost anything to it and the resulting expression is still 0. Then, we can add a term x to both side of the equation to reach an iteration form of x=g(x). Let us demonstrate this approach where we divide f(x) by 2x first, then add x.

Step 1. Reduce x^2 =a to the form f(x)=0

$$f(x)=x^2 - a=0$$

Step 2. Multiply by -1/2x (or some other expression of x)

$$\frac{x^2 - a}{-2 \cdot x} = \frac{0}{-2 \cdot x} = 0$$

Step 3. Add x to both sides (which is an easy way of generating the iteration form x=g(x)).

$$x=x-\frac{x^2-a}{2\cdot x}$$

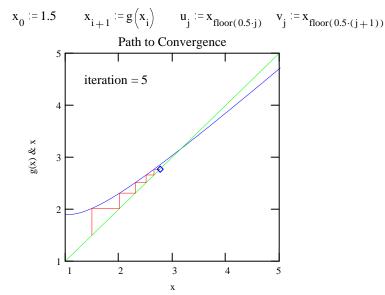
Step 4. Find g(x), which is the RHS of the last equation

$$g(x) := x - \frac{x^2 - a}{2 \cdot x}$$

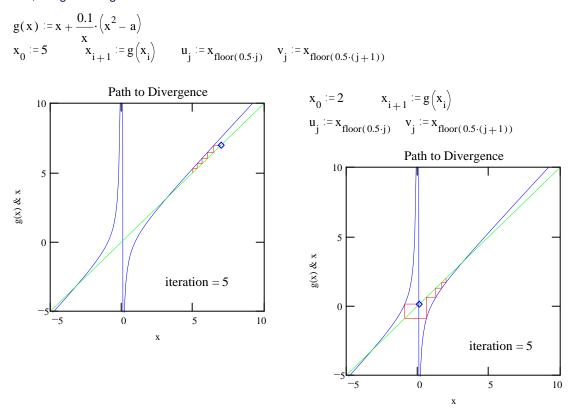
The resulting formula is the same as that from the Newton's method. Different numerical algorithms mainly differ in the expression mutiplied to f(x)=0 in Step 2. If we choose to multiply by -0.1/x instead, we get an expression that has a slower convergence property near the root. A slower convergence is not necessarily bad. We sometimes prefer a slower convergence when we want to approach the root gingerly and conservatively.

$$g(x) := x - \frac{0.1}{x} \cdot \left(x^2 - a\right)$$

Let us generate some numbers based on the above iteration formula and visualize convergence graphically.

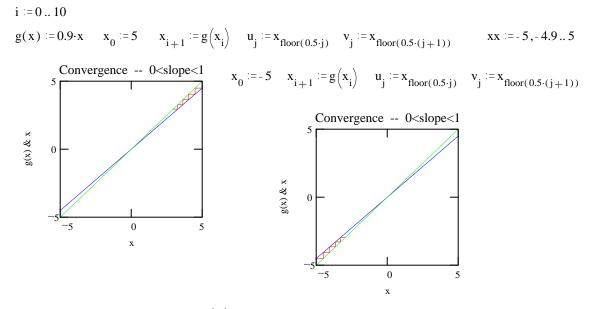


Of course, not all iteration formula lead to a root. On the other hand, if we choose to multiply by +0.1/x, we get divergence.

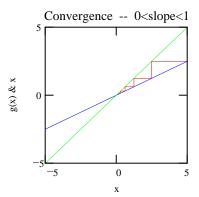


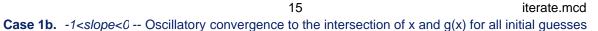
When do we achieve convergence and when do we face divergence? In general, we reach convergence when the |slope of g(x)| at the point of intersection with x is within -1 to 1. Furthermore, as shown below, we achieve very fast convergence when the slope of g(x) is close to 0. On the other hand, convergence is slow when the slope of g(x) is close to 1 or -1. Hence, the trick is to come up with an iteration scheme such that the slope of g(x) lies within -1 and 1, preferably close to 0. If we did not know anything about the convergence property, we have about 50% chance on the average of coming up with a converging formula (and 50% chance of a diverging one). So if the first try does not work, keep trying different manipulations to get x=g(x). The chances are you will hit one converging formula if you tried often enough. Below, we graphically demonstrate these two cases.

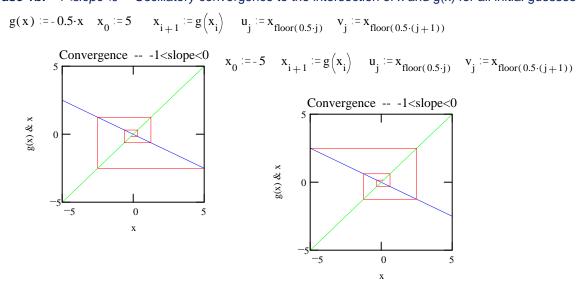
Case 1a. 0<slope<1 -- Monotonic convergence to the intersection of x and g(x) for all initial guesses



 $g(\mathbf{x}) \coloneqq 0.5 \cdot \mathbf{x} \qquad \mathbf{x}_0 \coloneqq 5 \qquad \mathbf{x}_{i+1} \coloneqq g(\mathbf{x}_i) \qquad \mathbf{u}_j \coloneqq \mathbf{x}_{\text{floor}(0.5 \cdot j)} \qquad \mathbf{v}_j \coloneqq \mathbf{x}_{\text{floor}(0.5 \cdot (j+1))}$







Case 2a. 1<slope -- Monotonic divergence from the intersection of x and g(x) for all initial guesses $g(x) := 1.5 \cdot x$ $x_0 := 0.5 \quad x_{i+1} := g(x_i) \quad u_j := x_{floor(0.5 \cdot j)} \quad v_j := x_{floor(0.5 \cdot (j+1))}$

