Accelerating fixed point iterations with Newton's method

Conor McCoid

May 30th, 2025

McMaster University

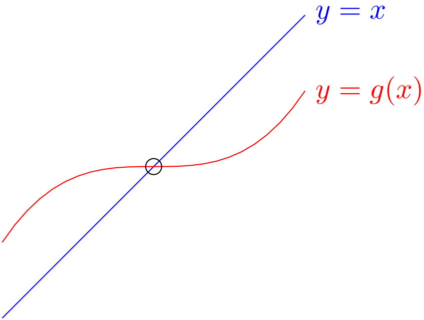
Fixed point iterations

Fixed points in 1D

The fixed point of a function $g(x)$ is a point x^* such that

$$g(x^*) = x^*.$$

These are also described as the intersection between the lines $y = g(x)$ and $y = x$.

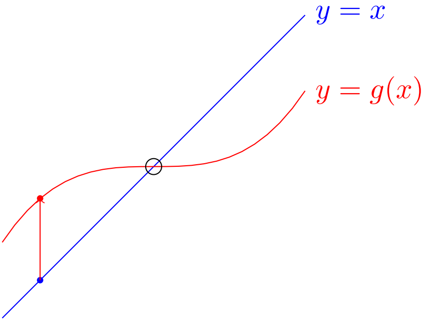


Fixed point iterations

To find fixed points of a given function $g(x)$, we can set up an iteration:

$$x_{n+1} = g(x_n).$$

This iteration stops at a fixed point. It converges if $g(x_n)$ is closer to x^* than x_n.

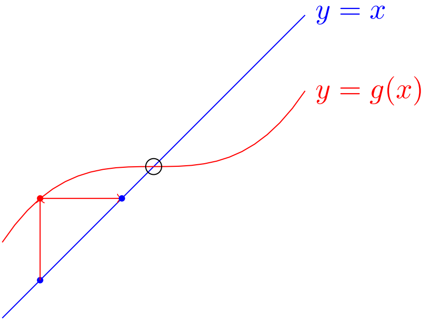


Fixed point iterations

To find fixed points of a given function $g(x)$, we can set up an iteration:

$$x_{n+1} = g(x_n).$$

This iteration stops at a fixed point. It converges if $g(x_n)$ is closer to x^* than x_n.

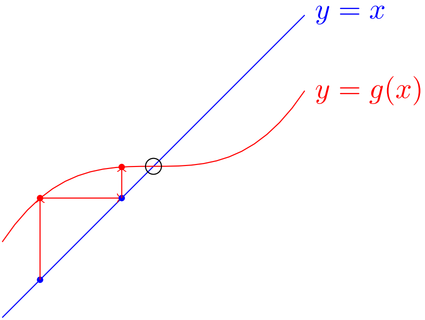


Fixed point iterations

To find fixed points of a given function $g(x)$, we can set up an iteration:

$$x_{n+1} = g(x_n).$$

This iteration stops at a fixed point. It converges if $g(x_n)$ is closer to x^* than x_n.

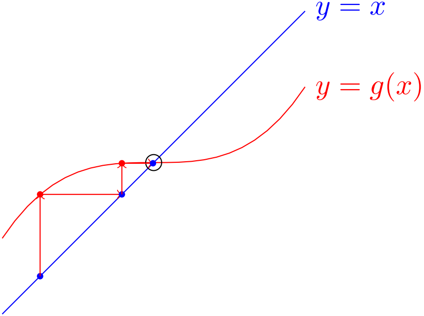


Fixed point iterations

To find fixed points of a given function $g(x)$, we can set up an iteration:

$$x_{n+1} = g(x_n).$$

This iteration stops at a fixed point. It converges if $g(x_n)$ is closer to x^* than x_n.



When does a fixed point iteration converge in 1D?

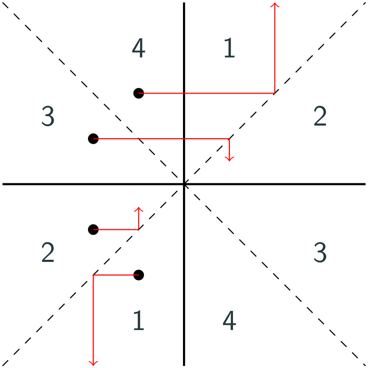
Convergence of the iteration $x_{n+1} = g(x_n)$ depends on which region $(x, g(x))$ lies.

1: Monotonic divergence

2: Monotonic convergence

3: Oscillatory convergence

4: Oscillatory divergence



Fixed point iterations in higher dimensions

Above 1D, fixed points of a vector-valued multivariate function $\boldsymbol{g}(\boldsymbol{x})$ satisfy

$$\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{x}^*.$$

Fixed point iterations are defined as

$$\boldsymbol{x}_{n+1} = \boldsymbol{g}(\boldsymbol{x}_n).$$

When does a fixed point iteration converge in nD?

Convergence can be shown with the Banach fixed-point theorem, which in this context requires

$$\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\| \leq q \|\boldsymbol{x} - \boldsymbol{y}\|, \quad q \in [0, 1).$$

In the 'regional' framework from 1D, we require

$$\|\boldsymbol{x}^* - \boldsymbol{g}(\boldsymbol{x}_n)\| < \|\boldsymbol{x}^* - \boldsymbol{x}_n\|.$$

This distinguishes between convergence and divergence, but monotonicity and oscillations are now harder to recognize.

Numerical methods as fixed point iterations

Almost every iterative method can be expressed as a fixed point iteration.

For example, consider Gauss-Seidel applied to:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix}$$

$$\boldsymbol{u}_{n+1} = A^{-1} \left(\boldsymbol{a} - B \boldsymbol{v}_n \right), \quad \boldsymbol{v}_{n+1} = D^{-1} \left(\boldsymbol{b} - C \boldsymbol{u}_{n+1} \right)$$

$$\implies \boldsymbol{v}_{n+1} = D^{-1} \Big(\boldsymbol{b} - C A^{-1} \left(\boldsymbol{a} - B \boldsymbol{v}_n \right) \Big) = \boldsymbol{g}(\boldsymbol{v}_n)$$

Continuous fixed point iterations

We can represent a fixed point iteration as the numerical integration of a function:

$$\frac{\boldsymbol{x}_{n+1} - \boldsymbol{x}_n}{\Delta t} = \boldsymbol{g}(\boldsymbol{x}_n) - \boldsymbol{x}_n, \quad \Delta t = 1,$$

$$\implies \frac{d\boldsymbol{x}}{dt} = \boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{x}.$$

Newton's method

Newton's method (in 1D)

Newton's method is used to find a root of a given function $f(x)$:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

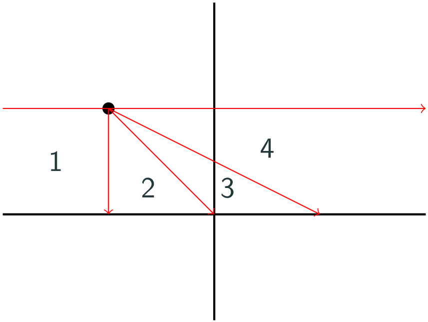
This can be viewed as a fixed point iteration:

$$g_f(x) = x - \frac{f(x)}{f'(x)},$$

then a fixed point of $g_f(x)$ is a root of $f(x)$.

When does Newton's method converge in 1D?

For Newton's method the regions of $g_f(x)$ depend on the slope of $f(x)$.
Most of the boundaries between the regions are known problems for Newton's method: a slope of zero divides regions 1 and 4, and an infinite slope divides 1 and 2.

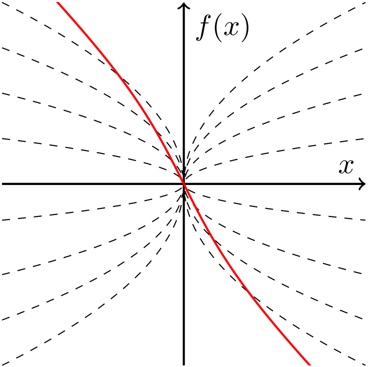


Cycles in Newton's method

The line between regions 3 and 4 can cause cycles in fixed point iterations. For Newton's method, this happens when $f(x)$ is parallel to

$$f_C(x) = C\sqrt{|x - x^*|}$$

for some constant $C \in \mathbb{R}$.



Continuous Newton's method

We can represent Newton's method as the numerical integration of an ODE:

$$\frac{x_{n+1} - x_n}{\Delta t} = -\frac{f(x_n)}{f'(x_n)}, \quad \Delta t = 1,$$

$$\implies \frac{\partial f}{\partial x}\frac{dx}{dt} = -f(x) \implies f(x(t)) = f(x(0))e^{-t}.$$

Newton's method in higher dimensions

Newton's method in higher dimensions requires the Jacobian of the function, $J_f(\boldsymbol{x})$:

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n - J_f^{-1}(\boldsymbol{x}_n)\boldsymbol{f}(\boldsymbol{x}_n).$$

The Kantorovich Theorem tells us this method converges as long as the initial guess is sufficiently close to the root (amongst other assumptions).

Davidenko-Branin trick

Davidenko (1953) and Branin (1972) suggest an update to Newton's method:

$$\frac{d\boldsymbol{x}}{dt} = \frac{\operatorname{adj} J_f}{|\det J_f|}\boldsymbol{f}(\boldsymbol{x}),$$

using some numerical integration scheme (the update is the absolute value around $\det J_f$).

Because $\det J_f$ only changes sign when passing over a root, this version of Newton's method will always travel in the same direction between roots. This allows the method to go over 'humps' in the function that would cause Newton to diverge otherwise.

Acceleration

Reposing a fixed point as a root

Given a function $\boldsymbol{g}(\boldsymbol{x})$ with a fixed point $\boldsymbol{x}^*$, we can make a function with a root:

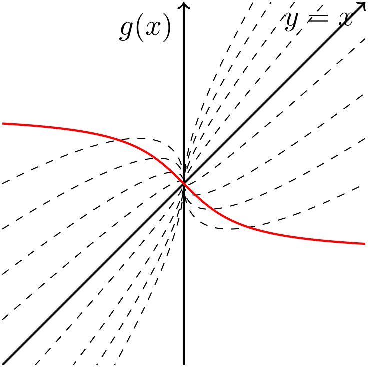
$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{x}.$$

There are an infinite number of ways to construct such a function, but this is the simplest.

Newton-accelerated FP in 1D

Apply the Newton analysis to the function $g(x)$. The boundaries of divergence are now when the slope of $g(x)$ is infinite, 1, or between 1 and parallel to

$$g_C(x) = C\sqrt{|x - x^*|} + x$$



When does Newton-accelerated FP converge in 1D?

We can show necessary and sufficient conditions for when accelerating by Newton will guarantee convergence, based on the behaviour of the iterative method.

$g(x)$ lies in	Necessary condition	Sufficient condition
1	$g'(x) > 1$	
2	$g'(x) < 1$	$g'(x) < 1/2$
3	$g'(x) < 1/2$	$g'(x) < 0$
4	$g'(x) < 0$	

Convergent iterative methods

It is sometimes possible to prove the guaranteed monotonic convergence of iterative methods, i.e. Schwarz methods with certain PDEs.

This means the fixed point for these methods is a contraction mapping, putting it in region 2 in the above framework.

A basic algorithm for 1D

Suppose $g(x)$ lies in region 2. Start with some initial guess x_0.

1. If $g'(x_n) = 1$, then accelerating with Newton will cause a division by infinity $\rightarrow$ use the fixed point iteration
2. If $|g'(x_n) - 1| < 1/2$, then using Newton with the Davidenko-Branin trick is guaranteed to be convergent $\rightarrow$ use Newton
3. Let $\tilde{x}$ be the point halfway between x_n and the Newton step; if the sign of $g(\tilde{x}) - \tilde{x}$ is the same as $g(x_n) - x_n$, then the fixed point lies between $\tilde{x}$ and the Newton step $\rightarrow$ use Newton
4. If none of these are true, use the fixed point iteration

What about higher dimensions?

This algorithm relies on the analysis of the boundary between regions 3 and 4. In 1D, that was lines parallel to $C\sqrt{|x - x^*|}$.

In higher dimensions, the boundary is significantly more complicated. It may not be possible to extract necessary and sufficient conditions from this analysis.

Augmenting Newton's method

If an iterative method is known to be convergent, then its fixed point iteration can help anchor Newton's method.

There are only three points of interest to an augmented Newton:

- the current iterate, $\boldsymbol{x}_n$;
- the fixed point step, $\boldsymbol{g}(\boldsymbol{x}_n)$;
- the Newton step, $\boldsymbol{F}(\boldsymbol{x}_n)$ (and possibly the Newton-Davidenko-Branin step).

Thus, we need only consider the 2D plane that contains these points. (The Newton, Newton-Davidenko-Branin, and current iterate all lie on the same line.)

Trust region

Since the method is convergent, $\boldsymbol{g}(\boldsymbol{x}_n)$ lies closer to $\boldsymbol{x}^*$ than $\boldsymbol{x}_n$.

Line search

The fixed point step must have a reasonable step size, unlike Newton which may leap a great distance away. But the Newton direction may be preferable.

Two-step

Take the fixed point step, then step towards the Newton step, but only part of the way.