A Subspace Acceleration Method for Fixed Point Iterations

Neil N. Carlson
Keith Miller

Fixed Point Iteration

Fixed point iteration for $x^\star = g(x^\star)$, $g : \mathbb{R}^m \to \mathbb{R}^m$:

x_0 given
for $n = 0, 1, 2, \ldots$ **do**
  $x_{n+1} = g(x_n)$
end for

$x_n \to x^\star$ if $\|x^\star - x_0\|$ and $\|\mathrm{D}g(x^\star)\|$ are sufficiently small.

Fixed point iteration for $f(x^\star) = 0$, $f : \mathbb{R}^m \to \mathbb{R}^m$, $(g(x) = x - f(x))$:

x_0 given
for $n = 0, 1, 2, \ldots$ **do**
  $x_{n+1} = x_n - f(x_n)$
end for

$x_n \to x^\star$ if $\|x^\star - x_0\|$ and $\|\mathrm{D}f(x^\star) - I\|$ are sufficiently small.

Modified Newton Methods: An Aside

Newton's method for $f(x^\star) = 0$, $f : \mathbb{R}^m \to \mathbb{R}^m$:

> x_0 given
> **for** $n = 0, 1, 2, \ldots$ **do**
> $\quad x_{n+1} = x_n - [\mathrm{D}f(x_n)]^{-1} f(x_n)$
> **end for**

Modified Newton's method replaces $\mathrm{D}f(x_n)$ with $P(x_n)$.
Example: $P(x_n) = \mathrm{D}f(\bar{x})$ (constant) for some $\bar{x}$.

Define $h(x) = [P(x)]^{-1} f(x)$. Then this modified Newton iteration is just a fixed point iteration for $h(x^\star) = 0$.

Accelerated FP Correction: Motivation

Rewrite our iteration as

> x_0 given
> **for** $n = 0, 1, 2, \ldots$ **do**
> $v_{n+1} = f(x_n)$ (Correction)
> $x_{n+1} = x_n - v_{n+1}$
> **end for**

If we were free to choose v_{n+1}, how would we choose it?
Perhaps as the solution of

$$0 = f(x_n - v_{n+1}) \approx f(x_n) - \mathrm{D}f(x_n)\, v_{n+1}.$$

FP iteration: Don't know $\mathrm{D}f(x_n)$, so just approximate it by I.

But if $\mathrm{D}f \approx$ constant, we **DO** know something about $\mathrm{D}f$!

The Accelerated FP Correction

To generate the correction v_{n+1} we have available:

$$\begin{aligned} \text{Corrections: } & v_1, \ldots, v_n, & V_n = [v_1 \cdots v_n], && \mathcal{V}_n = \operatorname{span}\{v_1, \ldots, v_n\} \\ f\text{-differences: } & w_1, \ldots, w_n, & W_n = [w_1 \cdots w_n], && \mathcal{W}_n = \operatorname{span}\{w_1, \ldots, w_n\} \end{aligned}$$

where $w_j = f(x_{j-1}) - f(x_j)$, ($w_j \approx \mathrm{D}f\, v_j$).

Idea: Split the correction $v_{n+1} = v' + v''$, with $v' \in \mathcal{V}_n$:

$$0 = f(x_n) - \mathrm{D}f(x_n)\,(v' + v'') \quad \rightsquigarrow \quad 0 = f(x_n) - \mathrm{D}f\, v' - I\, v''.$$

Accelerated correction (Carlson & Miller, SISC '98):

$$v_{n+1} = \underbrace{V_n z}_{\in \mathcal{V}_n} + \underbrace{\big(f(x_n) - W_n z\big)}_{\in \mathcal{W}_n^{\perp}}$$

where $z = \operatorname{argmin}_{\zeta \in \mathbb{R}^n} \| f(x_n) - W_n \zeta \|$.

The Nonlinear Reality

Of course $\mathrm{D}f$ isn't constant. In recognition of this fact we consider the most recent corrections and differences to be the most reliable.

- Use the w_j in reverse order.
- Only use a limited number of the most recent w_j.
- Drop any w_j that is nearly in the span of the preceding vectors.

The Accelerated FP Iteration

To summarize, the accelerated fixed point iteration is

$x := x_0$
repeat
  $v := \mathrm{FPA}\big(f(x)\big)$ ($v := f(x)$ is the unaccelerated correction)
  $x := x - v$
until converged

The acceleration procedure FPA is a black box. It sits in the loop listening to the sequence of function evaluations $f(x_0), f(x_1), \ldots$ and returning the accelerated corrections $v_1, v_2, \ldots$