Nonlinear Equations

Econ 5170

Computational Methods in Economics

2022-2023 Spring Term

Outline

1. One-Dimensional Problems
   - Bisection Method
   - Newton's Method
2. Multivariate Nonlinear Equations
   - Gauss-Jacobi and Gauss-Seidel Methods
   - Newton's Method
3. Global Convergence
   - Powell's Method
   - Continuation Method
   - Homotopy Method

General Problem

- Solve for a zero of a function

$$f(x) = 0$$

where $f : R^n \to R^n$.

- Special case: solve for a fixed point of a function:

$$f(x) = x \Leftrightarrow f(x) - x = 0$$

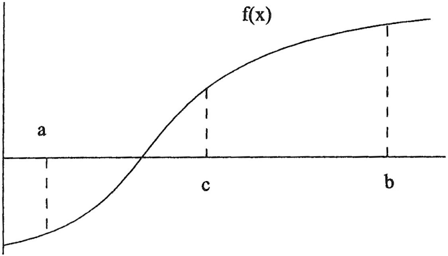
- Idea: generate a sequence of guesses that converges to the solution
- Univariate and multivariate problems

Bisection Method

- Univariate problem: $f : R \to R$
- Suppose f is continuous and $f(a) < 0 < f(b)$ for some $a, b, a < b$
- The intermediate value theorem tells that there is some zero of f in (a, b).
- Consider $c = \frac{1}{2}(a + b)$, the midpoint of $[a, b]$.
  - If $f(c) = 0$, we are done.
  - If $f(c) < 0$, there is a zero of f in (c, b). Continue by focusing on (c, b).
  - If $f(c) > 0$, there is a zero of f in (a, b). Continue by focusing on (a, c).

Algorithm

- Initialization: Initialize and bracket a zero: find $x^L < x^R$ such that $f(x^L)f(x^R) < 0$, and choose stopping rule parameters $\epsilon, \delta > 0$.
- Step 1. Compute midpoint: $x^M = (x^L + x^R)/2$
- Step 2. Refine the bounds: if $f(x^M)f(x^L) < 0$, $x^R = x^M$ and do not change x^L; else $x^L = x^M$ and leave x^R unchanged.
- Step 3. Check stopping rule: if $x^R - x^L \leq \epsilon(1 + |x^L| + |x^R|)$ or if $|f(x^M)| \leq \delta$, then stop and report solution at x^M; else go to step 1.



Bisection Method

Stopping rules

- Stop whenever the bracketing interval is so small that we do not care about any further precision: $x^R - x^L \leq \epsilon(1 + |x^L| + |x^R|)$
  - Change relative to x^L and x^R.
  - Avoid the problem where the solution is close to $x = 0$ and x^L and x^R converge to 0.
- Stop when $f(x^M)$ is less than the expected error in calculating f, which is controlled by δ.

Convergence

- Bisection method always converge to a solution once we have found an initial pair of points that bracket a zero and f is continuous.
- It's slow. It takes more than three iterations to add a decimal digit of accuracy.

Newton's Method

- Use smoothness properties of f to formulate a method that is fast when it works but may not always converge.
- Reduce a nonlinear problem to a sequence of linear problems
- Suppose the current guess is x_k. Construct the linear approximation to f at x_k

$$g(x) \equiv f'(x_k)(x - x_k) + f(x_k)$$

- Instead of solving for a zero of f, solve for a zero of g

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \tag{1}$$

- Convergence: Suppose that f is C^2 and that $f(x^*) = 0$. If x_0 is sufficiently close to x^*, $f'(x^*) \neq 0$, and $|f''(x^*)/f'(x^*)| < \infty$, the Newton sequence x_k defined by (1) converges to x^*, and it is quadratically convergent, that is,

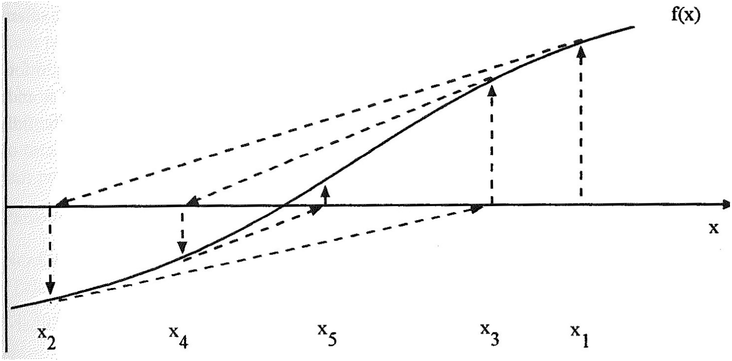
$$\lim_{k\to\infty} \sup \frac{|x_{k+1} - x^*|}{|x_k - x^*|^2} < \infty$$

Newton's Method

Algorithm

- Initialization. Choose stopping criterion ϵ and δ, and starting point x_0. Set $k = 0$.
- Step 1. Compute next iterate: $x_{k+1} = x_k - f(x_k)/f'(x_k)$.
- Step 2. Check stopping criterion: If $|x_k - x_{k+1}| \leq \epsilon(1 + |x_{k+1}|)$, go to step 3. Otherwise, go to step 1.
- Step 3. Report results and stop: If $|f(x_{k+1})| \leq \delta$, report success in finding a zero; otherwise, report failure.

Newton's Method



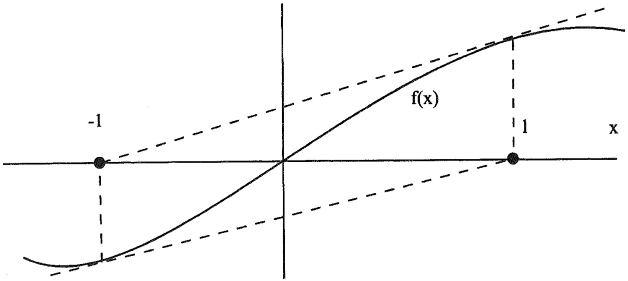
Newton's Method

Example 1:

$$f(x) = x^6$$

- $x_{k+1} = \frac{5}{6} x_k$
- Problem: x^6 is flat at its zero
- It's a slow, linearly convergent iteration
- Loose stopping rules may stop far from the true zero

Newton's Method



- $f'(1) = 0.5 = f'(-1)$ and $f(1) = 1 = f(-1)$.
- Converge to a cycle if starting at 1 or -1
- Converge if beginning with $x \in [-0.5, 0.5]$: importance of a good initial guess

Example 3: Competitive General Equilibrium

- Two goods and two consumers in an exchange economy.
- Agent $i, i = 1, 2$ has the utility function

$$u_i(x_1, x_2) = \frac{a_1^i x_1^{\eta_i+1}}{\eta_i + 1} + \frac{a_2^i x_2^{\eta_i+1}}{\eta_i + 1}$$

- If agent i has endowment $e^i \equiv (e_1^i, e_2^i)$ and the price of good j is p_j, then his demand function is

$$d_j^i(p) = \theta_j^i I^i p_j^{-\eta_i}$$

where $I^i = pe^i$ and $\theta_j^i = (a_j^i)^{\eta_i} / \sum_{l=1}^2 (a_l^i)^{\eta_i} p_l^{(1-\eta_i)}$

- An equilibrium solution is

$$\sum_{i=1}^2 d_1^i(p) = \sum_{i=1}^2 e_1^i, \quad p_1 + p_2 = 1$$

Newton's Method

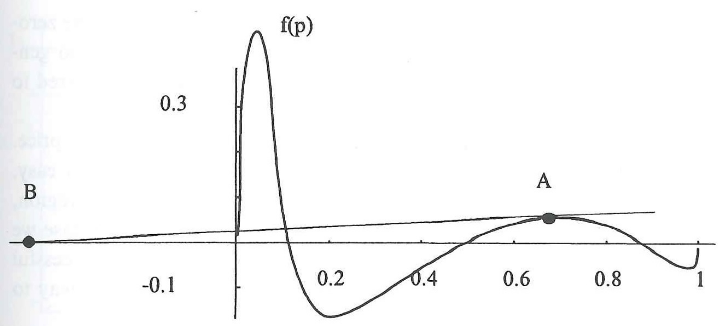
- $a_1^1 = a_2^2 = 1024, a_2^1 = a_1^2 = 1, e_1^1 = e_2^2 = 12, e_2^1 = e_1^2 = 1, \eta_1 = \eta_2 = -5$.
- Three equilibria:

$$p^1 = (0.5, 0.5), p^2 = (0.113, 0.887), p^3 = (0.887, 0.113)$$

- Reduce this problem to a one-variable problem by substituting $p_2 = 1 - p_1$.

$$f(p_1) = \sum_{i=1}^{2} d_1^i(p_1, 1 - p_1) - \sum_{i=1}^{2} e_1^i = 0$$

Newton's Method



Newton's Method

Limited domain problem: The excess demand function f is defined only for positive p_1.

- Solution 1: Check at each iteration whether $f(x^{k+1})$ is defined, and if it is not, move x^{k+1} toward x^k until f is defined.
  - Require access to the source code of the zero-finding routine

Newton's Method

- Solution 2: Extend the definition of f so that it is defined at any price.

$$\tilde{f}(p_1) = \begin{cases} f(p_1), & p_1 > \epsilon \\ f(\epsilon) + f'(\epsilon)(p_1 - \epsilon) + \dfrac{f''(\epsilon)(p_1 - \epsilon)^2}{2}, & p_1 \leq \epsilon \end{cases}$$

  - Replace f with a C^2 function that agrees with f at most positive prices and is defined for all prices.
  - Choose ϵ so that there are no solutions to $f(p_1) = 0$ in $(0, \epsilon)$.
  - When converge to a solution with a negative price, try again with a different initial guess

Newton's Method

- Solution 3: change the variable
  - Restate the problem in terms of $z \equiv P^{-1}(p_1)$ so $f(P(z)) = 0$.
  - If we want p_1 to stay within $[0, 1]$, use $p_1 = e^z/(e^z + e^{-z})$ with the inverse map $z = (1/2)\ln(p_1/(1 - p_1))$.
  - Newton's method applied to

$$g(z) \equiv f(P(z)) = 0$$

results in the iteration

$$z_{k+1} = z_k - \frac{g(z_k)}{g'(z_k)}$$

Secant Method

- A key step in Newton's method is the computation of $f'(x)$, which may be costly
- The secant method employs the idea of linear approximations but never evaluates f'
- It approximates $f'(x_k)$ with the slope of the secant of f between x_k and x_{k-1}

$$x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}$$

- If $f(x^*) = 0, f'(x^*) \neq 0$, and $f'(x)$ and $f''(x)$ are continuous near x^*, the secant method converges at the rate $(1+\sqrt{5})/2$.

Solve a Non-linear Equation

Use three algorithms to solve the following non-linear equation:

$$1 + x + \log x = 0$$

- Bisection method: choose two initial values that bound a zero
- Newton's method: try different initial guesses
- Secant method: try different initial guesses

Multivariate Nonlinear Equations

- $f : R^n \to R^n$ and we solve $f(x) = 0$. n equations in n unknowns

$$
\begin{aligned}
f^1(x_1, x_2, \ldots, x_n) &= 0 \\
&\vdots \\
f^n(x_1, x_2, \ldots, x_n) &= 0
\end{aligned}
$$

Gauss-Jacobi Method

$$
\begin{aligned}
f^1(x_1^{k+1}, x_2^k, \ldots, x_n^k) &= 0 \\
f^2(x_1^k, x_2^{k+1}, \ldots, x_n^k) &= 0 \\
&\vdots \\
f^n(x_1^k, x_2^k, \ldots, x_n^{k+1}) &= 0
\end{aligned}
$$

- There are $n!$ different Gauss-Jacobi schemes. If some equation depends on only one unknown, then that equation should be equation 1 and that variable should be variable 1.
- Linear Gauss-Jacobi method: takes a single Newton step to approximate the solution to the nonlinear equation

$$
x_i^{k+1} = x_i^k - \frac{f^i(x^k)}{f^i_{x_i}(x^k)}, \quad i = 1, .., n
$$

Gauss-Seidel Method

$$
\begin{aligned}
f^1(x_1^{k+1}, x_2^k, \ldots, x_n^k) &= 0 \\
f^2(x_1^{k+1}, x_2^{k+1}, \ldots, x_n^k) &= 0 \\
&\vdots \\
f^n(x_1^{k+1}, x_2^{k+1}, \ldots, x_n^{k+1}) &= 0
\end{aligned}
$$

Linear Gauss-Seidel method:

$$x_i^{k+1} = x_i^k - \frac{f^i}{f^i_{x_i}}(x_1^{k+1}, ..., x_{i-1}^{k+1}, x_i^k, ..., x_n^k), \quad i = 1, .., n$$

Gauss-Jacobi and Gauss-Seidel Method

- We can apply stabilization and acceleration methods to attain or accelerate convergence just as with linear equations.
- However, convergence is at best linear.
- For $x^{k+1} = G(x^k)$, the spectral radius of the Jacobian evaluated at the solution, $G_x(x^*)$, is its asymptotic linear rate of convergence.
- Stopping rule: stop when $||x^{k+1} - x^k|| \leq (1-\beta)\epsilon$ where $\beta = \rho(G_x(x^*))$. Computing β directly would be impractical. Estimate β with

$$\tilde{\beta} = \max\left\{ \frac{||x^{k-j+1} - x^k||}{||x^{k-j} - x^k||} \quad j = 2, .., L \right\}$$

  for some L. The estimate $\tilde{\beta}$ would be close to $\rho(G_x(x^*))$ if $x^k \approx x^*$.
- Accept x^{k+1} if $||f(x^{k+1})|| < \delta$.

Gauss-Jacobi and Gauss-Seidel Method

Example: Duopoly

- Two goods, Y and Z, and the utility function over those goods and money M is
$$U(Y,Z) = (1 + Y^\alpha + Z^\alpha)^{\eta/\alpha} + M$$
with $\alpha = 0.999$ and $\eta = 0.2$. Unit cost of Y is 0.07 and of Z is 0.08.
- Profit for the Y firm is $\Pi^Y(Y,Z) = Y(p_Y(Y,Z) - 0.07)$ where p_Y is the price of Y.
Profit for the Z firm is $\Pi^Z(Y,Z) = Z(p_Z(Y,Z) - 0.08)$ where p_Z is the price of Z.
- We solve the system
$$\Pi_1^Y(Y,Z) = 0$$
$$\Pi_2^Z(Y,Z) = 0$$

Gauss-Jacobi and Gauss-Seidel Method

Example: Duopoly

- Keep Y and Z positive, so instead solve

$$\Pi_1^Y(e^y, e^z) = 0$$
$$\Pi_2^Z(e^y, e^z) = 0$$

  for $y = \ln Y$ and $z = \ln Z$.

- There is a unique Cournot-Nash equilibrium at the intersection of the reaction curves, $(y^*, z^*) = (-0.137, -0.576)$, or $(Y^*, Z^*) = (0.87, 0.56)$.

Gauss-Jacobi and Gauss-Seidel Method

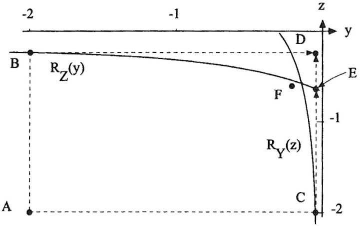


Figure 5.7
Solving the duopoly problem (5.4.7)

Table 5.1
Errors of Gaussian methods applied to (5.4.7)

Iteration	Gauss-Jacobi	Linear Gauss-Jacobi	Gauss-Seidel	Linear Gauss-Seidel
1	$(1(-1), 3(-1))$	$(1(0), 1(0))$	$(1(-1), -6(-2))$	$(1(0), -8(0))$
2	$(-7(-2), -6(-2))$	$(-8(-1), -1(-1))$	$(1(-3), -6(-3))$	$(-6(-1), 4(2))$
3	$(1(-2), 4(-2))$	$(2(-1), 1(-1))$	$(1(-4), -6(-4))$	*
4	$(-7(-3), -6(-3))$	$(-8(-1), -2(-1))$	$(1(-5), -6(-5))$	*
5	$(1(-3), 4(-3))$	$(1(-1), 3(-1))$	$(9(-6), -5(-6))$	*
6	$(-7(-4), -5(-4))$	$(-5(-2), -6(-2))$	$(9(-7), -5(-7))$	*
7	$(9(-5), 4(-4))$	$(9(-3), 3(-2))$	$(8(-8), -5(-8))$	*
8	$(-6(-5), -5(-5))$	$(-5(-3), -5(-3))$	$(8(-9), -5(-9))$	*

Note: The * means that the iterates became infinite.

Newton's Method

- By Taylor's theorem, the linear approximation of f around the initial guess x^0 is

$$f(x) \doteq f(x^0) + J(x^0)(x - x^0)$$

  The Newton iteration scheme is

$$x^{k+1} = x^k - J(x^k)^{-1} f(x^k)$$

- Newton's method is quadratically convergent. The critical assumption is that $det(J(x^*)) \neq 0$.

Broyden Method (Secant)

- Explicit computation of Jacobian is often costly to compute and code.
- One would typically use finite differences to compute $J(x)$ in Newton's method but that requires n^2 evaluations of f.
- Broyden method begins with a rough guess of the Jacobian and use the successive evaluations of f and its gradient to update the guess of J.
    - Suppose the guess for the Jacobian at x^k is A_k. Use A_k to compute the Newton step, that is, solve $A_k s^k = -f(x^k)$ and define $x^{k+1} = x^k + s^k$.
    - Choose A_{k+1} consistent with the secant equation $f(x^{k+1}) - f(x^k) = A_{k+1} s^k$.
    - For any direction q orthogonal to s^k we have no information about $f(x^{k+1}) - f(x^k)$. Assume that $A_{k+1} q = A_k q$ whenever $q^\top s^k = 0$.

$$A_{k+1} = A_k + \frac{(f(x^{k+1}) - f(x^k) - A_k s^k)(s^k)^\top}{(s^k)^\top s^k}$$

- Broyden method converges superlinearly.

Newton's Method vs. Broyden methods

Table 5.3
Errors of Newton and Broyden methods applied to (5.4.7)

Iterate k	Newton	Broyden
0	$(-0.19(1), -0.14(1))$	$(-0.19(1), -0.14(1))$
1	$(0.55(0), 0.28(0))$	$(0.55(0), 0.28(0))$
2	$(-0.59(-1), 0.93(-2))$	$(0.14(-1), 0.65(-2))$
3	$(0.15(-3), 0.81(-3))$	$(-0.19(-2), 0.40(-3))$
4	$(0.86(-8), 0.54(-7))$	$(0.45(-3), 0.24(-3))$
5	$(0.80(-15), 0.44(-15))$	$(-0.11(-3), 0.61(-4))$
6	$(0, 0)$	$(0.26(-4), -0.14(-4))$
7	$(0, 0)$	$(-0.60(-5), 0.33(-5))$
8	$(0, 0)$	$(0.14(-5), -0.76(-6))$
9	$(0, 0)$	$(-0.32(-6), 0.18(-6))$
10	$(0, 0)$	$(0.75(-7), 0.41(-7))$

Optimization and Nonlinear Equations

- None of the above methods is globally convergent.
- Optimization $\Rightarrow$ nonlinear equations:
  If $f(x)$ is C^2, the solution to $\min_x f(x)$ is also a solution to the system of first-order conditions $\nabla f(x) = 0$.
- Nonlinear equations $\Rightarrow$ optimization:
  Any solution to the system $f(x) = 0$ is also a global solution to

$$\min_x \sum_{i=1}^{n} f^i(x)^2 \tag{2}$$

Problems:

  - There may be local minima that are not near any solution to $f(x) = 0$
  - The condition number of the Hessian of (2) is roughly the square of the condition number of the Jacobian of $f(x) = 0$.

Powell's Hybrid Method

- Newton's method converges quickly if it converges but it may diverge.
- The minimization idea of (2) will converge to something, but it may do so slowly.

Powell's Hybrid Method

- Define $SSR(x) = \sum_{i=1}^{n} f^i(x)^2$. It indicates how well we are doing and help restrain Newton's method.
- Check if a Newton's step reduces the value of SSR. If yes, accept $x^k + s^k$
- Otherwise, choose a direction equal to a combination of the Newton step and the gradient of $-SSR$
  Matlab: check steepest descent
- This method may stop if they come too near a local minimum of SSR, and we can continue by choosing a new starting point.

Overidentified System

- Solve a nonlinear system $f(x) = 0$ where $f : R^n \to R^m$ and $n < m$.
- Solve the least square problem

$$\min_x f(x)^\top f(x)$$

- Can be solved using the optimization methods.

A Simple Continuation Method

- Solve $f(x;t) = 0$, where $f : R^n \times R \rightarrow R^n$, for some specific value $t = t^*$.
- x is a list of endogenous variables and t is a parameter of taste, technology, or policy.
- Suppose we do not have a solution for $f(x;t^*) = 0$ but we do know that for t^0, $f(x;t^0) = 0$ has a solution x^0.
- If t^* is near t^0, x^0 will be a good initial guess when trying to solve $f(x;t^*) = 0$.
- An appropriate initial guess for the Jacobian for the t^* problem is the last approximation for the Jacobian of the t^0 problem.
- If t^* is not close to t^0, we can still construct a sequence of problems of the form $f(x;t) = 0$ satisfying

$$t^0 \approx t^1 \approx t^2 \approx \ldots \approx t^n \approx t^*$$

A Simple Continuation Method

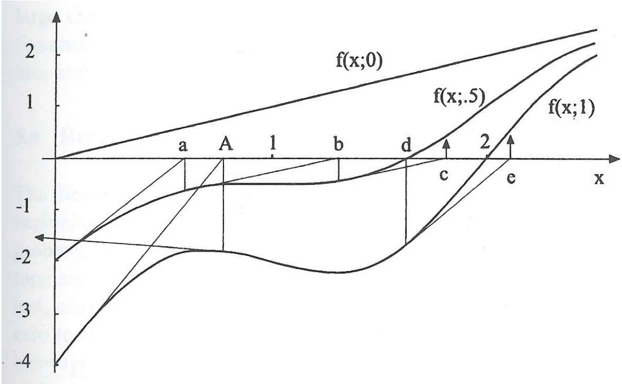
Algorithm

- Initialization. Form the sequence $t^0 \approx t^1 \approx t^2 \approx \ldots \approx t^n \approx t^*$; set $i = 0$.
- Step 1. Solve $f(x; t^{i+1}) = 0$ using x^i as the initial guess; set x^{i+1} equal to the solution.
- Step 2. If $i + 1 = n$, report x^n as the solution to $f(x; t^*)$ and stop; else go to step 1.

A Simple Continuation Method

Example 1:

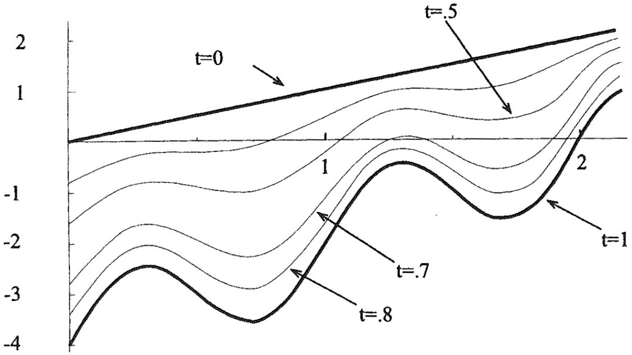
$$f(x;t) = (1-t)x + t(2x - 4 + \sin(\pi x))$$



Newton's method alone starting at $x = 0$ would fail because of the oscillations in $\sin(\pi x)$.

Example 2:

$$H(x;t) = (1-t)x + t(2x - 4 + \sin(2\pi x))$$



A Simple Continuation Method

Problem of the simple continuation method:

- $H(x; 0.70)$ has three zeros on $[0, 2]$.
- $H(x; 0.74) = 0$ has a solution $x = 1.29$. But $H(x; 0.75) = 0$ has a solution $x = 1.92$. Using $i = 1.29$ as the initial guess for $H(x; 0.75) = 0$ does not lead to convergence.
- $x = 1.29$ is close to two zeros of $H(x, 0.74)$ but is not close to any zero of $H(x; 0.75) = 0$

Homotopy Method

- Homotopy method: A globally convergent way to find zeros of $f : R^n \to R^n$.
- A homotopy function $H : R^{n+1} \to R^n$ that continuously deform g into f is any continuous function H where

$$H(x, 0) = g(x), \quad H(x, 1) = f(x)$$

- In practice, $H(x, 0)$ is a simple function with easily calculated zeros, and $H(x, 1)$ is the function whose zeros we want.
  - Newton homotopy: $H(x, t) = f(x) - (1 - t)f(x^0)$ for some x^0. At $t = 0$, $H = f(x) - f(x^0)$ which has a zero at $x = x^0$.
  - Fixed-point homotopy: $H(x, t) = (1 - t)(x - x^0) + tf(x)$ for some x^0.
  - Linear homotopy: $H(x, t) = tf(x) + (1 - t)g(x)$.

Homotopy Method

- The basic object is the set

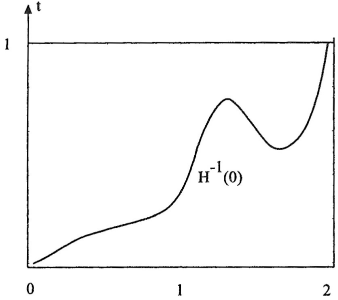
$$H^{-1}(0) = \{(x,t) | H(x,t) = 0\}$$

If $H(x,0)$ and $H(x,1)$ have zeros, the hope is that there is a continuous path in $H^{-1}(0)$ which connects zeros of $H(x,0)$ to zeros of $H(x,1)$.

Homotopy Method

Example 2:

$$H(x, t) = (1 - t)x + t(2x - 4 + \sin(2\pi x))$$

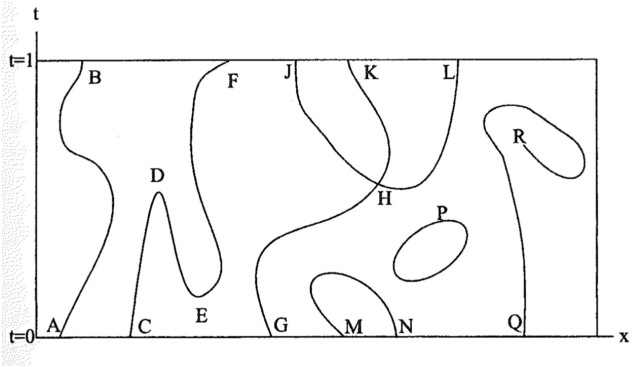


- At $t = 0$ and $t = 1$, there are unique zeros.
- For $t \in (0.53, 0.74)$, there are three zeros.

Homotopy Method

- Simple continuation assumes that we can proceed from the zero of the $t = 0$ problem to the zero of the $t = 1$ problem by taking an increasing sequence of t values.
- Homotopy methods instead follow the path $H^{-1}(0)$, tracing it wherever it goes and allowing t to increase and decrease as necessary to stay on $H^{-1}(0)$.

Homotopy Method



Homotopy Method

Parametric path following

- Parameterize both x and t in terms of a third parameter s. The parametric path satisfies $H(x(s), t(s)) = 0$ for all s.

$$\sum_{i=1}^{n} H_{x_i}(x(s), t(s))x_i'(s) + H_t(x(s), t(s))t'(s) = 0$$

- Define $y(s) = (x(s), t(s))$, then y obeys the system of differential equations

$$\frac{dy_i}{ds} = (-1)^i det\left(\frac{\partial H}{\partial y}(y)_{-i}\right), \quad i = 1, ..., n+1$$

where $(\cdot)_{-i}$ means we remove the ith column.

Homotopy Method

Example 2:

$$\begin{pmatrix} dx/ds \\ dt/ds \end{pmatrix} = \begin{pmatrix} -H_t \\ H_x \end{pmatrix} = \begin{pmatrix} x - (2x - 4 + \sin(2\pi x)) \\ 1 - t + t(2 + 2\pi\cos(2\pi x)) \end{pmatrix}$$

- To find a zero of $H(x, 1)$, we start with $(x, t) = (0, 0)$ and then solve the above differential equations until we reach $t(s) = 1$.
- Define

$$x_{i+1} = x_i + h(x_i - (2x_i - 4 + \sin(2\pi x_i)))$$
$$t_{i+1} = t_i + h(1 - t_i + t_i(2 + 2\pi\cos(2\pi x_i)))$$

where $h = 0.001$ is the step size corresponding to ds.

Homotopy Method

Table 5.5
Homotopy path following in (5.9.6)

Iterate i	t_i	x_i	True solution $H(x, t_i) = 0$
50	0.05621	0.16759	0.16678
100	0.11130	0.30748	0.30676
200	0.15498	0.64027	0.64025
300	0.35543	1.02062	1.01850
400	0.69540	1.23677	1.23288
500	0.67465	1.41982	1.42215
600	0.51866	1.68870	1.69290
650	0.61731	1.84792	1.84391
700	0.90216	1.97705	1.97407
711	0.99647	2.00203	1.99915
712	1.00473	2.00402	2.00114