2021 Spring MAS 365
Chapter 10: Numerical Solutions of Nonlinear Systems of Equations

Donghwan Kim

KAIST

Apr/8,13, 2021

1. 10.1 Fixed Points for Functions of Several Variables
2. 10.2 Newton's Method
3. 10.4 Steepest Descent Techniques

System of Nonlinear Equations

- A system of nonlnear equations has the form

$$\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= 0, \\ f_2(x_1, x_2, \ldots, x_n) &= 0, \\ \vdots \qquad\qquad & \quad \vdots \\ f_n(x_1, x_2, \ldots, x_n) &= 0. \end{aligned}$$

- Using vector notation, this can be simply written as

$$\boldsymbol{F}(\boldsymbol{x}) = \mathbf{0},$$

where $f_1, f_2, \ldots, f_n$ are called the **coordinate functions** of $\boldsymbol{F} = (f_1, f_2, \ldots, f_n)^t$.

Functions from $\mathbb{R}^n$ into $\mathbb{R}$

Definition 1

Let f be a function defined on a set $D \subset \mathbb{R}^n$ and mapping into $\mathbb{R}$. The function f is said to have the **limit** *L at $\boldsymbol{x}_0$, written*

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = L,$$

if, given any number $\epsilon > 0$, a number $\delta > 0$ exists with

$$|f(\boldsymbol{x}) - L| < \epsilon$$

whenever $\boldsymbol{x} \in D$, and

$$0 < ||\boldsymbol{x} - \boldsymbol{x}_0|| < \delta.$$

any norm

Functions from $\mathbb{R}^n$ into $\mathbb{R}$ (cont'd)

Definition 2

Let f be a function from a set $D \subset \mathbb{R}^n$ into $\mathbb{R}$. The function f is **continuous** *at $\boldsymbol{x}_0 \in D$ provided $\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x})$ exists and*

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = f(\boldsymbol{x}_0).$$

Moreover, f is **continuous** *on a set D if f is continuous at every point of D. This concept is expressed by writing $f \in C(D)$.*

Functions from $\mathbb{R}^n$ into $\mathbb{R}^n$

Definition 3

Let $\boldsymbol{F}$ be a function from $D \subset \mathbb{R}^n$ into $\mathbb{R}^n$ of the form

$$\boldsymbol{F}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_n(\boldsymbol{x}))^t,$$

where f_i is mapping from $\mathbb{R}^n$ into $\mathbb{R}$ for each i. We define

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{L} = (L_1, L_2, \ldots, L_n)^t,$$

if and only if $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} f_i(\boldsymbol{x}) = L_i$, for each $i = 1, 2, \ldots, n$.

- The function $\boldsymbol{F}$ is **continuous** at $\boldsymbol{x}_0 \in D$ provided $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \boldsymbol{F}(\boldsymbol{x})$ exists and $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{F}(\boldsymbol{x}_0)$. In addition, $\boldsymbol{F}$ is continuous on the set D if $\boldsymbol{F}$ is continuous at each $\boldsymbol{x}$ in D. This concept is expressed by writing $\boldsymbol{F} \in C(D)$.

Fixed Points in $\mathbb{R}^n$

Definition 4

A function $\boldsymbol{G}$ from $D \subset \mathbb{R}^n$ into $\mathbb{R}^n$ has a **fixed point** *at $\boldsymbol{p} \in D$ if $\boldsymbol{G}(\boldsymbol{p}) = \boldsymbol{p}$.*

Fixed Points in $\mathbb{R}^n$ (cont'd)

Theorem 1

Let $D = \{(x_1, x_2, \ldots, x_n)^t \mid a_i \le x_i \le b_i,$ *for each* $i = 1, 2, \ldots, n\}$ *for some collection of constants* $a_1, a_2, \ldots, a_n$ *and* $b_1, b_2, \ldots, b_n$*. Suppose* $\boldsymbol{G}$ *is a continuous function from* $D \subset \mathbb{R}^n$ *into* $\mathbb{R}^n$ *with the property that* $\boldsymbol{G}(\boldsymbol{x}) \in D$ *whenever* $\boldsymbol{x} \in D$*. Then* $\boldsymbol{G}$ *has a fixed point in* D*.*

Moreover, suppose that all the component functions of $\boldsymbol{G}$ *have continuous partial derivatives and a constant* $K < 1$ *exists with*

$$\left|\frac{\partial g_i(\boldsymbol{x})}{\partial x_j}\right| \le \frac{K}{n}, \quad \text{whenever } \boldsymbol{x} \in D,$$

for each $j = 1, 2, \ldots, n$ *and each component function* g_i*. Then the fixed-point sequence* $\{\boldsymbol{x}^{(k)}\}_{k=0}^{\infty}$ *defined by an arbitrarily selected* $\boldsymbol{x}^{(0)}$ *in* D *and generated by*

$$\boldsymbol{x}^{(k)} = \boldsymbol{G}(\boldsymbol{x}^{(k-1)}), \quad \text{for each } k \ge 1$$

converges to the unique fixed point $\boldsymbol{p} \in D$ *and*

$$||\boldsymbol{x}^{(k)} - \boldsymbol{p}||_\infty \le \frac{K^k}{1-K}||\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}||_\infty.$$

Linear rate

Fixed Points in $\mathbb{R}^n$ (cont'd)

Ex Find a fixed-point form of the nonlinear system

$$\begin{aligned} 3x_1 - \cos(x_2 x_3) - \frac{1}{2} &= 0, \\ x_1^2 - 81(x_2 + 0.1)^2 + \sin x_3 + 1.06 &= 0, \\ e^{-x_1 x_2} + 20x_3 + \frac{10\pi - 3}{3} &= 0. \end{aligned}$$

1. 10.1 Fixed Points for Functions of Several Variables
2. 10.2 Newton's Method
3. 10.4 Steepest Descent Techniques

Recall: Newton's Method in One Variable

Theorem 2

Let p be a solution of the equation $x = g(x)$. Suppose that $g'(p) = 0$ and g'' is continuous with $|g''(x)| < M$ on an open interval I containing p.

Then there exists a $\delta > 0$ such that, for $p_0 \in [p - \delta, p + \delta]$, the sequence defined by $p_n = g(p_{n-1})$, when $n \geq 1$, converges at least **quadratically** *to p.*

Moreover, for sufficiently large values of n,

$$|p_{n+1} - p| < \frac{M}{2} |p_n - p|^2.$$

- We found a function ϕ with the property that

$$g(x) = x - \phi(x) f(x)$$

gives qudaratic convergence to the fixed point p of the function g.

Newton's Method

Theorem 3

Let $\boldsymbol{p}$ be a solution of $\boldsymbol{G}(\boldsymbol{x}) = \boldsymbol{x}$. suppose a number $\delta > 0$ exists with the properties:

- *$\partial g_i / \partial x_j$ is continuous on $N_\delta = \{\boldsymbol{x} \mid ||\boldsymbol{x} - \boldsymbol{p}|| < \delta\}$, for each $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n$;*
- *$\partial^2 g_i(\boldsymbol{x}) / (\partial x_j \partial x_k)$ is continuous, and $|\partial^2 g_i(\boldsymbol{x}) / (\partial x_j \partial x_k)| \leq M$ for some constant M, whenever $\boldsymbol{x} \in N_\delta$, for each $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n$, and $k = 1, 2, \ldots, n$;*
- *$\partial g_i(\boldsymbol{p}) / \partial x_k = 0$, for each $i = 1, 2, \ldots, n$, and $k = 1, 2, \ldots, n$.*

Then a number $\hat{\delta} \leq \delta$ exists such that the sequence generated by $\boldsymbol{x}^{(k)} = \boldsymbol{G}(\boldsymbol{x}^{(k-1)}$ converges quadratically to $\boldsymbol{p}$ for any choice of $\boldsymbol{x}^{(0)}$, provided that $||\boldsymbol{x}^{(0)} - \boldsymbol{p}|| < \hat{\delta}$. Moreover,

$$||\boldsymbol{x}^{(k)} - \boldsymbol{p}||_\infty \leq \frac{n^2 M}{2} ||\boldsymbol{x}^{(k-1)} - \boldsymbol{p}||_\infty^2, \quad \text{for each } k \geq 1.$$

Newton's Method (cont'd)

- Consider

$$\boldsymbol{G}(\boldsymbol{x}) = \boldsymbol{x} - A(\boldsymbol{x})^{-1}\boldsymbol{F}(x),$$

  where $A(\boldsymbol{x}) = [a_{ij}(\boldsymbol{x})]$.

- Assume that $A(\boldsymbol{x})$ is nonsingular near a solution $\boldsymbol{p}$ of $\boldsymbol{F}(\boldsymbol{x}) = \mathbf{0}$ and let $A(\boldsymbol{x})^{-1} = [b_{ij}(\boldsymbol{x})]$.

Newton's Method (cont'd)

- Since $g_i(\boldsymbol{x}) = x_i - \sum_{j=1}^{n} b_{ij}(\boldsymbol{x}) f_j(\boldsymbol{x})$, we have

$$\frac{\partial g_i}{\partial x_k}(\boldsymbol{x}) = \begin{cases} 1 - \sum_{j=1}^{n} \left(b_{ij}(\boldsymbol{x}) \frac{\partial f_j}{\partial x_k}(\boldsymbol{x}) + \frac{\partial b_{ij}}{\partial x_k}(\boldsymbol{x}) f_j(\boldsymbol{x}) \right), & i = k, \\ - \sum_{j=1}^{n} \left(b_{ij}(\boldsymbol{x}) \frac{\partial f_j}{\partial x_k}(\boldsymbol{x}) + \frac{\partial b_{ij}}{\partial x_k}(\boldsymbol{x}) f_j(\boldsymbol{x}) \right), & i \neq k, \end{cases}$$

Jacobian Matrix

- Define the matrix $J(\boldsymbol{x})$ by

$$J(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f_1}{\partial x_2}(\boldsymbol{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\boldsymbol{x}) \\ \frac{\partial f_2}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f_2}{\partial x_2}(\boldsymbol{x}) & \cdots & \frac{\partial f_2}{\partial x_n}(\boldsymbol{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f_n}{\partial x_2}(\boldsymbol{x}) & \cdots & \frac{\partial f_n}{\partial x_n}(\boldsymbol{x}) \end{bmatrix}$$

Newton's Method

Ex Find a fixed-point form of the nonlinear system

$$
\begin{aligned}
3x_1 - \cos(x_2 x_3) - \frac{1}{2} &= 0, \\
x_1^2 - 81(x_2 + 0.1)^2 + \sin x_3 + 1.06 &= 0, \\
e^{-x_1 x_2} + 20x_3 + \frac{10\pi - 3}{3} &= 0.
\end{aligned}
$$

1 10.1 Fixed Points for Functions of Several Variables

2 10.2 Newton's Method

3 10.4 Steepest Descent Techniques

Minimization and System of Nonlinear Equations

- A system of nonlinear equations

$$\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= 0, \\ f_2(x_1, x_2, \ldots, x_n) &= 0, \\ \vdots \qquad &\quad \vdots \\ f_n(x_1, x_2, \ldots, x_n) &= 0, \end{aligned}$$

has a solution $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^t$ when the function g defined by

$$g(x_1, x_2, \ldots, x_n) =$$

has the minimal value 0.

The Gradient of a Function

$V_k = -\nabla g(x_k)$

- For $g : \mathbb{R}^n \to \mathbb{R}$, the **gradient** of g at $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^t$ is denoted $\nabla g(\boldsymbol{x})$ and defined by

$$\nabla g(\boldsymbol{x}) = \left(\frac{\partial g}{\partial x_1}(\boldsymbol{x}), \frac{\partial g}{\partial x_2}(\boldsymbol{x}), \ldots, \frac{\partial g}{\partial x_n}(\boldsymbol{x}) \right)^t$$

The Gradient of a Function (cont'd)

steepest descent

- The **directional derivative** of g at $\boldsymbol{x}$ in the direction of $\boldsymbol{v}$ is defined by

$$D_{\boldsymbol{v}} g(\boldsymbol{x}) = \lim_{h \to 0} \frac{1}{h}[g(\boldsymbol{x} + h\boldsymbol{v}) - g(\boldsymbol{x})] = \boldsymbol{v}^t \nabla g(\boldsymbol{x}).$$

- This measures the change in the value of the function g relative to the change in the variable in the direction of $\boldsymbol{v}$.

The Gradient of a Function (cont'd)

$\exists \varepsilon > 0,\ g(x + hv) < g(x)$ for $\forall h \in (0, \varepsilon)$

$D_v g(x) = v^T \nabla g(x)$

$\min_{\|v\|=1} D_v g(x)$

- Steepest descent method updates as

$$\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} - \alpha \nabla g(\boldsymbol{x}^{(0)}),$$

for some constant $\alpha > 0$.

- Consider

$$h(\alpha) = g(\boldsymbol{x}^{(0)} - \alpha \nabla g(\boldsymbol{x}^{(0)}))$$

Steepest Descent Method

Ex Find a fixed-point form of the nonlinear system

$$\begin{aligned} 3x_1 - \cos(x_2x_3) - \frac{1}{2} &= 0, \\ x_1^2 - 81(x_2+0.1)^2 + \sin x_3 + 1.06 &= 0, \\ e^{-x_1x_2} + 20x_3 + \frac{10\pi-3}{3} &= 0. \end{aligned}$$