UNIVERSITY OF WATERLOO
FACULTY OF ENGINEERING
Department of Electrical &
Computer Engineering

ECE 204 *Numerical methods*

Newton's method in *n* dimensions

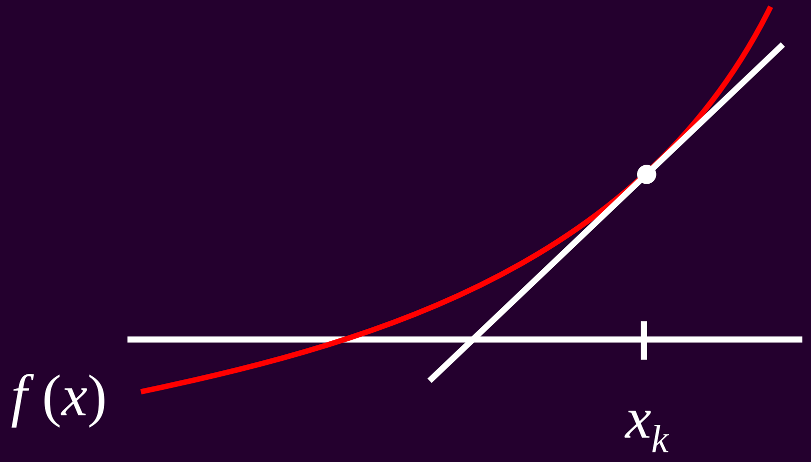
Douglas Wilhelm Harder, LEL, M.Math.
dwharder@uwaterloo.ca
dwharder@gmail.com

Introduction

- In this topic, we will
  - Derive Newton's method in *n* dimensions
  - Observe the formula is analogous to the one we've seen
  - Look at an example
  - Observe the rate of convergence is still $O(h^2)$

Newton’s method

- Recall Newton’s method:
  - Find the root of the tangent line at $(x_k, f(x_k))$



Generalization of Newton's method

- Recall the Taylor series for a real-valued function of a vector variable

$$f(\mathbf{u}) \approx f(\mathbf{u}_0) + \vec{\nabla} f(\mathbf{u}_0) \cdot (\mathbf{u} - \mathbf{u}_0)$$

$$= f(\mathbf{u}_0) + \begin{pmatrix} \frac{\partial}{\partial u_1} f(\mathbf{u}_0) & \cdots & \frac{\partial}{\partial u_n} f(\mathbf{u}_0) \end{pmatrix} (\mathbf{u} - \mathbf{u}_0)$$

- Suppose we have *n* such functions, so for each we have:

$$f_k(\mathbf{u}) \approx f_k(\mathbf{u}_0) + \begin{pmatrix} \frac{\partial}{\partial u_1} f_k(\mathbf{u}_0) & \cdots & \frac{\partial}{\partial u_n} f_k(\mathbf{u}_0) \end{pmatrix} (\mathbf{u} - \mathbf{u}_0)$$

Generalization of Newton's method

- Thus, we have:

$$f_1(\mathbf{u}) \approx f_1(\mathbf{u}_0) + \begin{pmatrix} \frac{\partial}{\partial u_1} f_1(\mathbf{u}_0) & \cdots & \frac{\partial}{\partial u_n} f_1(\mathbf{u}_0) \end{pmatrix} (\mathbf{u} - \mathbf{u}_0)$$

$$f_2(\mathbf{u}) \approx f_2(\mathbf{u}_0) + \begin{pmatrix} \frac{\partial}{\partial u_1} f_2(\mathbf{u}_0) & \cdots & \frac{\partial}{\partial u_n} f_2(\mathbf{u}_0) \end{pmatrix} (\mathbf{u} - \mathbf{u}_0)$$

$$\vdots$$

$$f_n(\mathbf{u}) \approx f_n(\mathbf{u}_0) + \begin{pmatrix} \frac{\partial}{\partial u_1} f_n(\mathbf{u}_0) & \cdots & \frac{\partial}{\partial u_n} f_n(\mathbf{u}_0) \end{pmatrix} (\mathbf{u} - \mathbf{u}_0)$$

$$\begin{pmatrix} f_1(\mathbf{u}) \\ f_2(\mathbf{u}) \\ \vdots \\ f_n(\mathbf{u}) \end{pmatrix} \qquad \begin{pmatrix} f_1(\mathbf{u}_0) \\ f_2(\mathbf{u}_0) \\ \vdots \\ f_n(\mathbf{u}_0) \end{pmatrix} \qquad \begin{pmatrix} \frac{\partial}{\partial u_1} f_1(\mathbf{u}_0) & \frac{\partial}{\partial u_2} f_1(\mathbf{u}_0) & \cdots & \frac{\partial}{\partial u_n} f_1(\mathbf{u}_0) \\ \frac{\partial}{\partial u_1} f_2(\mathbf{u}_0) & \frac{\partial}{\partial u_2} f_2(\mathbf{u}_0) & \cdots & \frac{\partial}{\partial u_n} f_2(\mathbf{u}_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial u_1} f_n(\mathbf{u}_0) & \frac{\partial}{\partial u_2} f_n(\mathbf{u}_0) & \cdots & \frac{\partial}{\partial u_n} f_n(\mathbf{u}_0) \end{pmatrix} (\mathbf{u} - \mathbf{u}_0)$$

Generalization of Newton's method

- Thus, we have:

$\mathbf{f}(\mathbf{u}_0)$

$\mathbf{f}(\mathbf{u})$

$$\begin{pmatrix} f_1(\mathbf{u}) \\ \vdots \\ f_n(\mathbf{u}) \end{pmatrix} \approx \begin{pmatrix} f_1(\mathbf{u}_0) \\ \vdots \\ f_n(\mathbf{u}_0) \end{pmatrix} + \begin{pmatrix} \frac{\partial}{\partial u_1} f_k(\mathbf{u}_0) & \cdots & \frac{\partial}{\partial u_n} f_1(\mathbf{u}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial u_1} f_n(\mathbf{u}_0) & \cdots & \frac{\partial}{\partial u_n} f_n(\mathbf{u}_0) \end{pmatrix} (\mathbf{u} - \mathbf{u}_0)$$

This is the Jacobian evaluated at $\mathbf{u}_0$

$$J(\mathbf{f})(\mathbf{u}) = \begin{pmatrix} \frac{\partial}{\partial u_1} f_k(\mathbf{u}) & \cdots & \frac{\partial}{\partial u_n} f_1(\mathbf{u}) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial u_1} f_n(\mathbf{u}) & \cdots & \frac{\partial}{\partial u_n} f_n(\mathbf{u}) \end{pmatrix}$$

Generalization of Newton's method

- Thus, we have n tangent $(n-1)$-dimensional hyperplanes at $\mathbf{u}_0$:

$$\mathbf{f}(\mathbf{u}) \approx \mathbf{f}(\mathbf{u}_0) + J(\mathbf{f})(\mathbf{u}_0)(\mathbf{u}-\mathbf{u}_0)$$

- A root of these hyperplanes may be found by equating this to the zero vector:

$$\mathbf{0} = \mathbf{f}(\mathbf{u}_0) + J(\mathbf{f})(\mathbf{u}_0)(\mathbf{u}-\mathbf{u}_0)$$

$$J(\mathbf{f})(\mathbf{u}_0)(\mathbf{u}-\mathbf{u}_0) = -\mathbf{f}(\mathbf{u}_0)$$

- This is a system of n linear equations in n unknowns
  - Let $\Delta\mathbf{u}_0 = \mathbf{u} - \mathbf{u}_0$, so we are solving

$$J(\mathbf{f})(\mathbf{u}_0)\Delta\mathbf{u}_0 = -\mathbf{f}(\mathbf{u}_0)$$

  - Having found $\Delta\mathbf{u}_0$, we now assign $\mathbf{u}_1 \leftarrow \mathbf{u}_0 + \Delta\mathbf{u}_0$
- We can now repeat this until $\left\|\mathbf{u}_{k+1}-\mathbf{u}_k\right\|_2 < \varepsilon_{\text{step}}$ and $\left\|\mathbf{f}(\mathbf{u}_{k+1})\right\|_2 < \varepsilon_{\text{abs}}$

Generalization of Newton's method

- You may be wondering, how are these related?

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

Solve $J(\mathbf{f})(\mathbf{u}_k)\Delta\mathbf{u}_k = -\mathbf{f}(\mathbf{u}_k)$ for $\Delta\mathbf{u}_k$,
and assign $\mathbf{u}_{k+1} \leftarrow \mathbf{u}_k + \Delta\mathbf{u}_k$

$$0 = f(x_k) + f^{(1)}(x_k)(x_{k+1} - x_k)$$

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

$$f^{(1)}(x_k)\Delta x_k = -f(x_k) \qquad \Delta x_k = -\frac{f(x_k)}{f^{(1)}(x_k)}$$

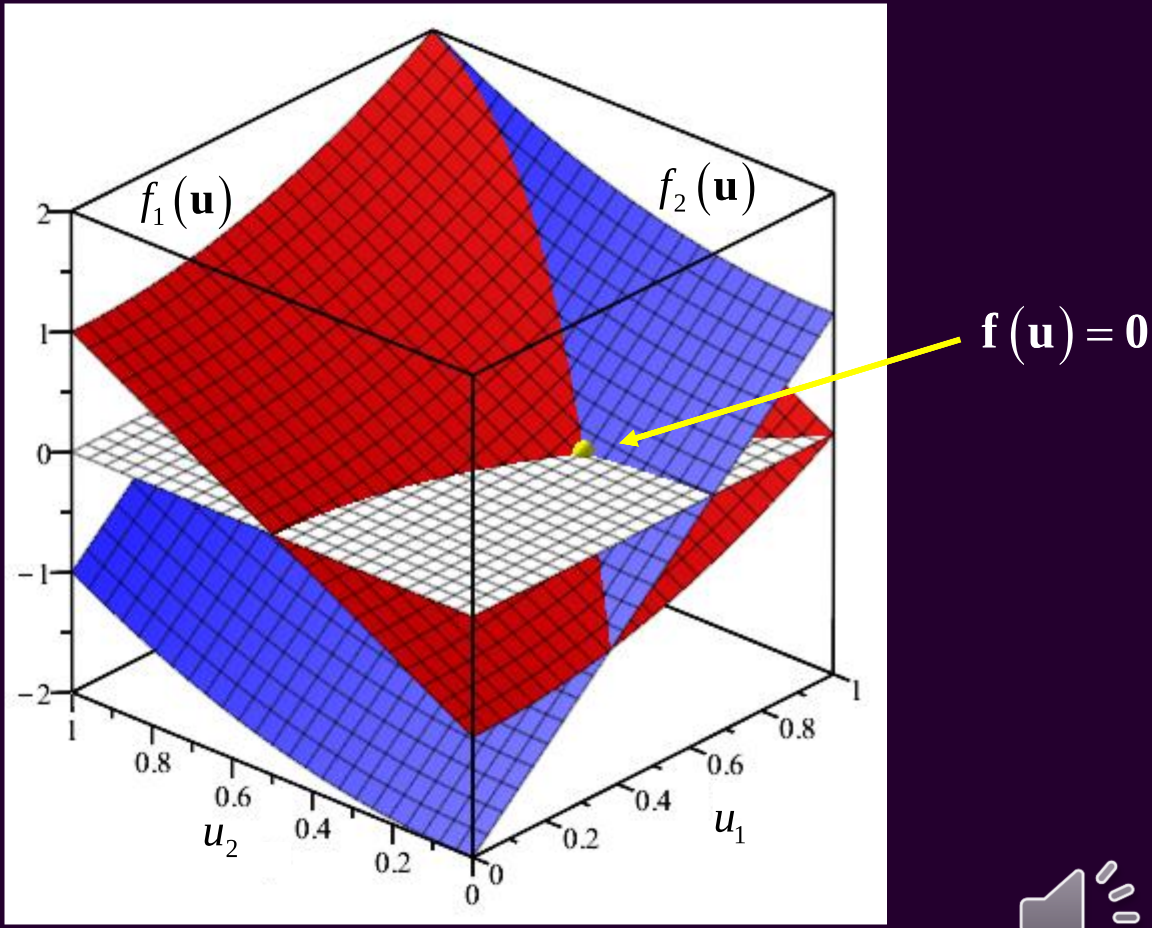
$$x_{k+1} = x_k + \Delta x_k$$

That is, $x_{k+1} = x_k - \frac{f(x_k)}{f^{(1)}(x_k)}$

Example

- Suppose we have the following:

$$\mathbf{f}(\mathbf{u}) = \begin{pmatrix} u_1^2 + 2u_2 - 1 \\ u_2^2 + 3u_1 - 2 \end{pmatrix}$$

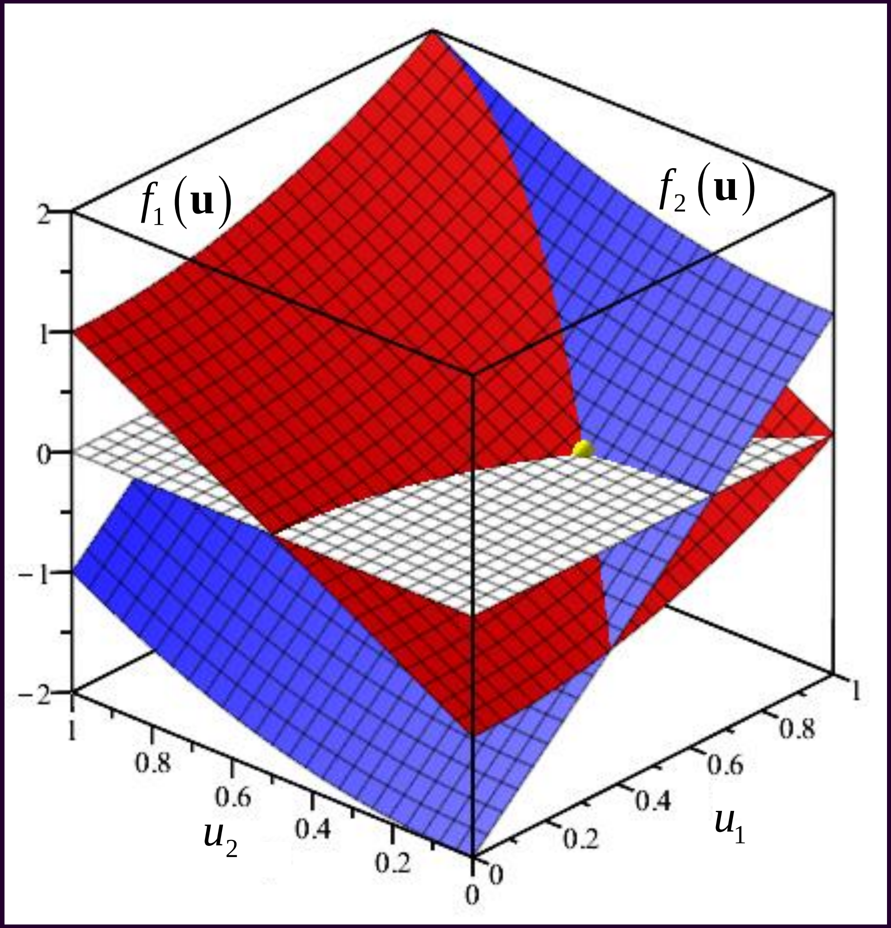


Example

- First, we calculate the Jacobian

$$\mathbf{f}(\mathbf{u}) = \begin{pmatrix} u_1^2 + 2u_2 - 1 \\ u_2^2 + 3u_1 - 2 \end{pmatrix}$$

$$J(\mathbf{f})(\mathbf{u}) = \begin{pmatrix} 2u_1 & 2 \\ 3 & 2u_2 \end{pmatrix}$$

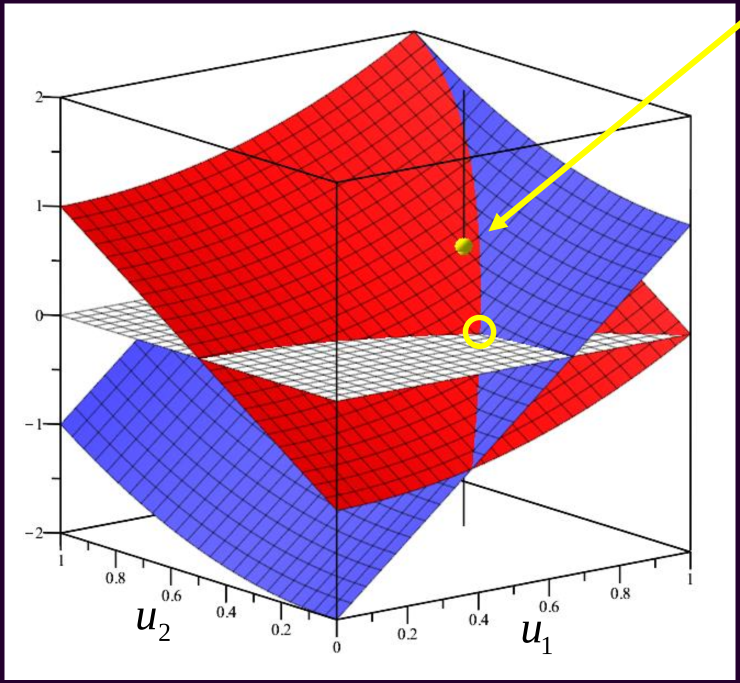


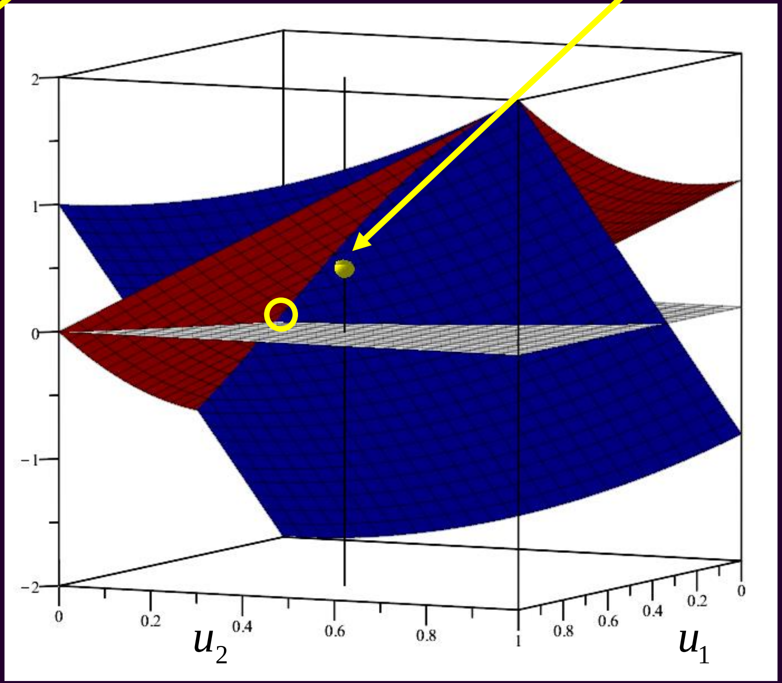
Example

- Note that $\mathbf{f}\begin{pmatrix}0.75\\0.5\end{pmatrix}=\begin{pmatrix}0.5625\\0.5\end{pmatrix}$
  - Thus, $\mathbf{u}_0=\begin{pmatrix}0.75\\0.5\end{pmatrix}$ is sort-of close to a root

$f_1(\mathbf{u}_0)=0.5625$

$f_2(\mathbf{u}_0)=0.5$



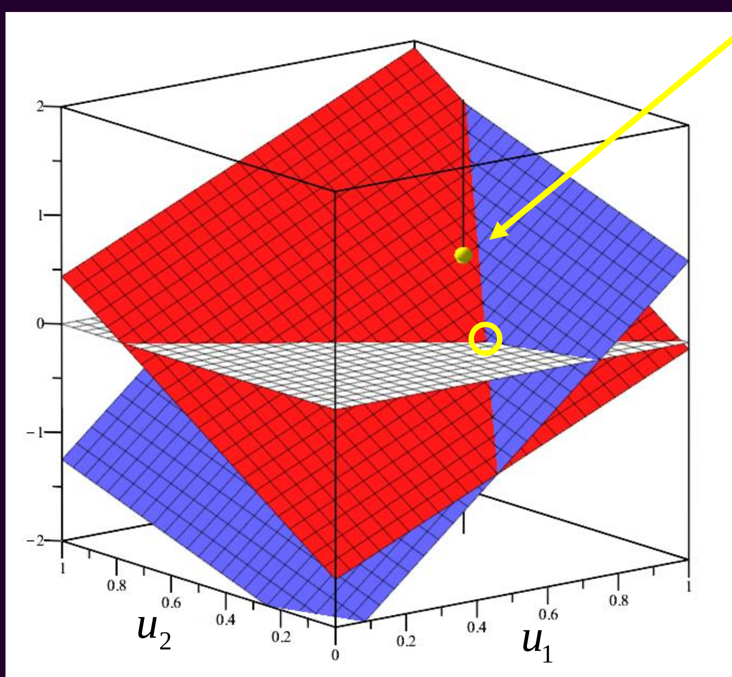


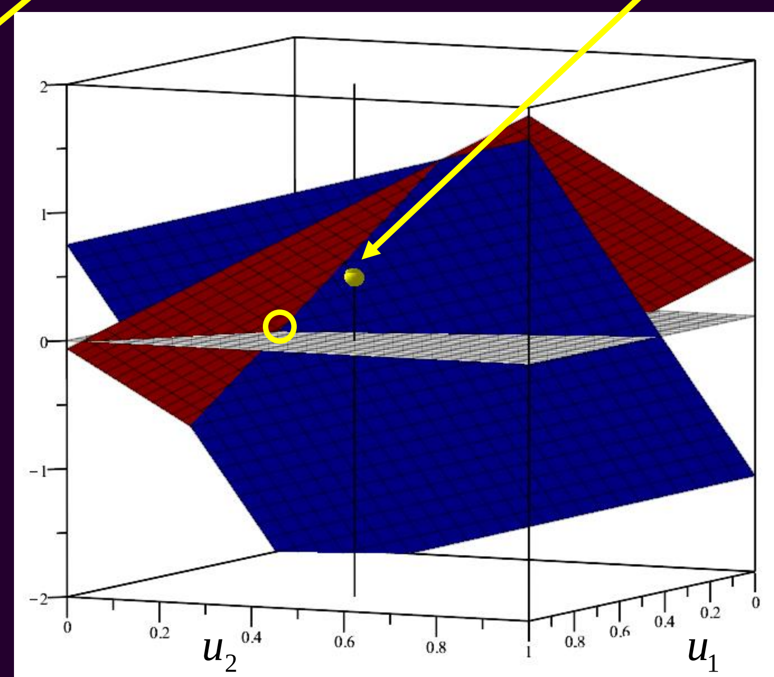
Example

- We can find tangent planes at each of these points

$$J(\mathbf{f})(\mathbf{u}_0)\Delta\mathbf{u}_0 = -\mathbf{f}(\mathbf{u}_0) \qquad \begin{pmatrix} 1.5 & 2 \\ 3 & 1 \end{pmatrix}\Delta\mathbf{u}_0 = \begin{pmatrix} -0.5625 \\ -0.5 \end{pmatrix} \qquad \Delta\mathbf{u}_0 = \begin{pmatrix} -0.0972222 \\ -0.2083333 \end{pmatrix}$$

$f_1(\mathbf{u}_0) = 0.5625$ $\qquad$ $f_2(\mathbf{u}_0) = 0.5$

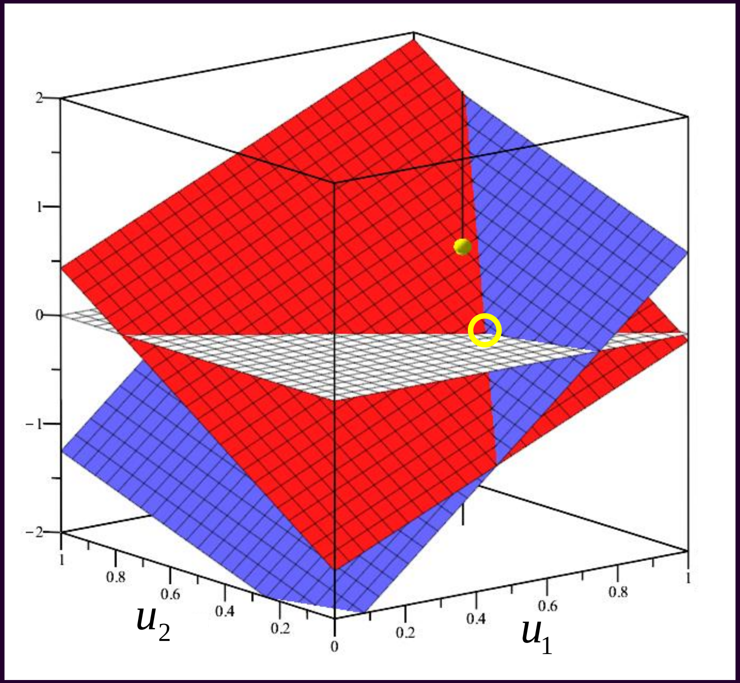


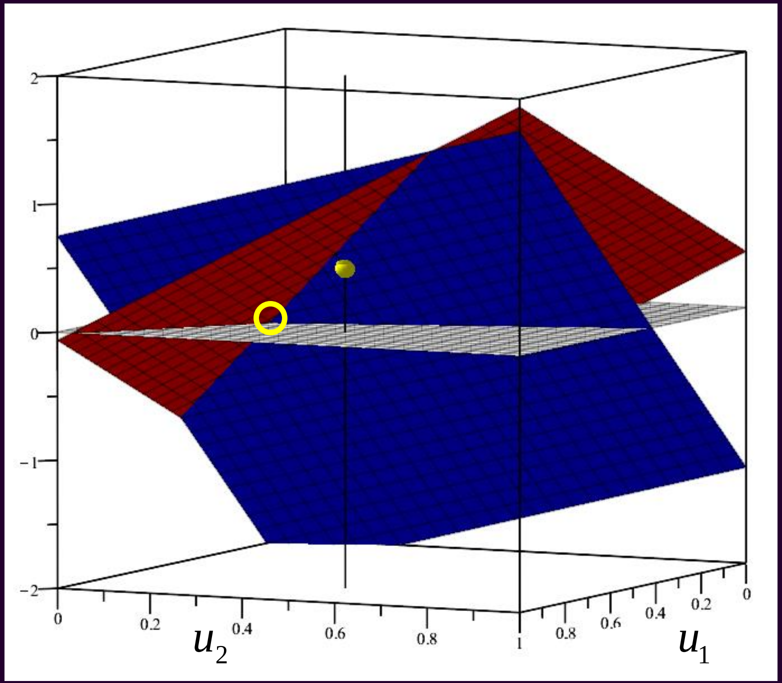


$$\mathbf{u}_0 = \begin{pmatrix} 0.75 \\ 0.5 \end{pmatrix}$$

Example

- We solved for $\Delta\mathbf{u}_0$ and so $\mathbf{u}_1 \leftarrow \mathbf{u}_0 + \Delta\mathbf{u}_0$
  - Thus, $\mathbf{u}_1 = \begin{pmatrix} 0.6527778 \\ 0.2916667 \end{pmatrix}$ and $\mathbf{f}\left(\mathbf{u}_1\right) = \begin{pmatrix} 0.009452 \\ 0.04340 \end{pmatrix}$





Example

- Here is a sequence of iterations:

$$\mathbf{u}_0 = \begin{pmatrix} 0.75 \\ 0.5 \end{pmatrix} \qquad \mathbf{f}(\mathbf{u}_0) = \begin{pmatrix} 0.5625 \\ 0.5 \end{pmatrix}$$

$$\mathbf{u}_1 = \begin{pmatrix} 0.6527777777777778 \\ 0.2916666666666667 \end{pmatrix} \qquad \mathbf{f}(\mathbf{u}_1) = \begin{pmatrix} 0.009452 \\ 0.04340 \end{pmatrix}$$

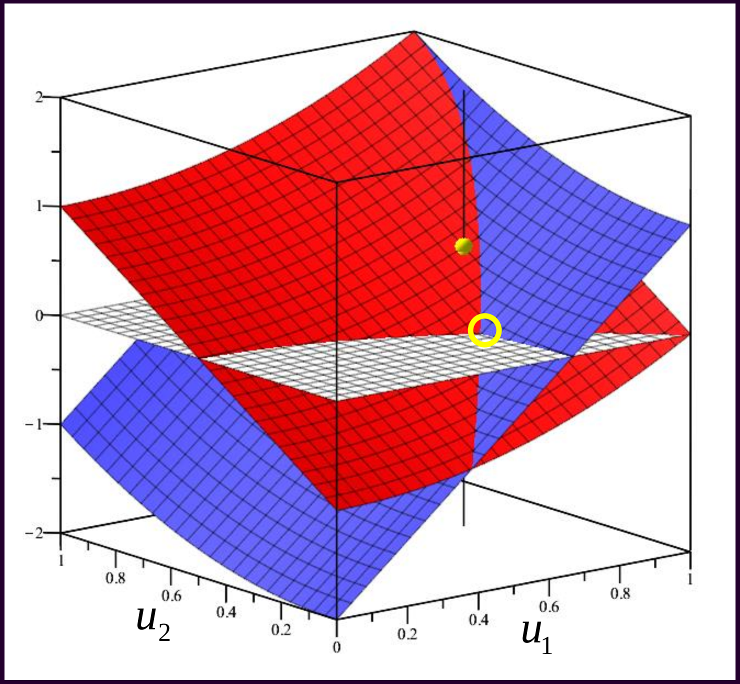
$$\mathbf{u}_2 = \begin{pmatrix} 0.6372594147395296 \\ 0.2970706289586095 \end{pmatrix} \qquad \mathbf{f}(\mathbf{u}_2) = \begin{pmatrix} 0.0002408 \\ 0.00002920 \end{pmatrix}$$

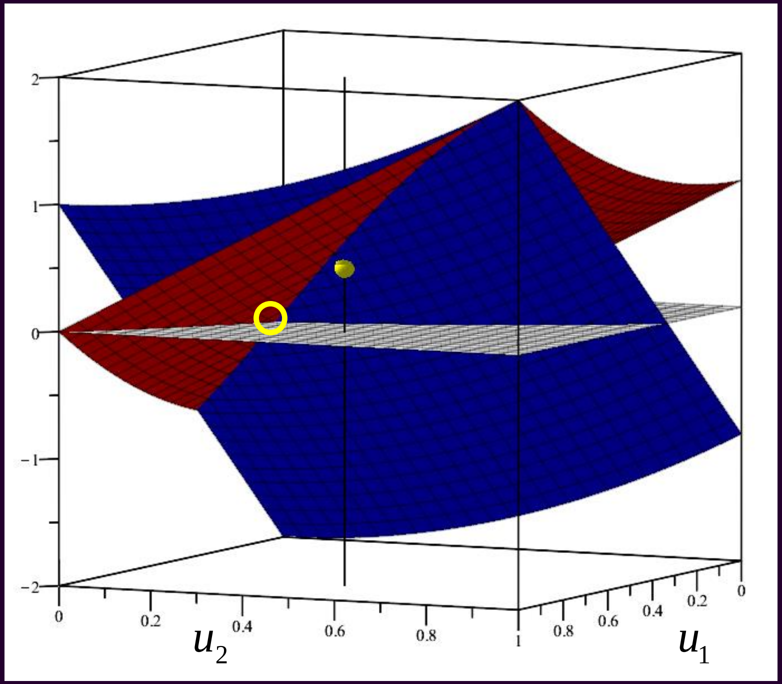
$$\mathbf{u}_3 = \begin{pmatrix} 0.6372755656421493 \\ 0.2969399268481651 \end{pmatrix} \qquad \mathbf{f}(\mathbf{u}_3) = \begin{pmatrix} 0.000000002609 \\ 0.0000001708 \end{pmatrix}$$

Example

- The actual root is closer to

$$\mathbf{u} = \begin{pmatrix} 0.6372755591552685 \\ 0.2969399308516699 \end{pmatrix} \qquad \mathbf{u}_3 = \begin{pmatrix} 0.6372755656421493 \\ 0.2969399268481651 \end{pmatrix}$$





Summary

- Following this topic, you now
  - Understand the generalization of Newton’s method
    - In two dimensions, we have two expressions in two variables
    - Given an initial approximation, find two tangent planes, and find the simultaneous root of those tangent planes
  - Know this generalizes to *n* dimensions
    - Find the tangent hyper-planes and find the root of the tangent hyper-planes
  - Are aware that the convergence is still $O(h^2)$

References

[1] https://en.wikipedia.org/wiki/Newton%27s_method

Acknowledgments

Jeffrey Cornelis for noting I left out the most significant digit in the leading entry of the approximation

$$\mathbf{u}_1 = \begin{pmatrix} 0.6527778 \\ 0.2916667 \end{pmatrix}$$

Colophon

These slides were prepared using the Cambria typeface. Mathematical equations use Times New Roman, and source code is presented using `Consolas`. Mathematical equations are prepared in MathType by Design Science, Inc.

Examples may be formulated and checked using Maple by Maplesoft, Inc.

The photographs of flowers and a monarch butter appearing on the title slide and accenting the top of each other slide were taken at the Royal Botanical Gardens in October of 2017 by Douglas Wilhelm Harder. Please see

https://www.rbg.ca/

for more information.

Disclaimer

These slides are provided for the ECE 204 *Numerical methods* course taught at the University of Waterloo. The material in it reflects the author's best judgment in light of the information available to them at the time of preparation. Any reliance on these course slides by any party for any other purpose are the responsibility of such parties. The authors accept no responsibility for damages, if any, suffered by any party as a result of decisions made or actions based on these course slides for any other purpose than that for which it was intended.