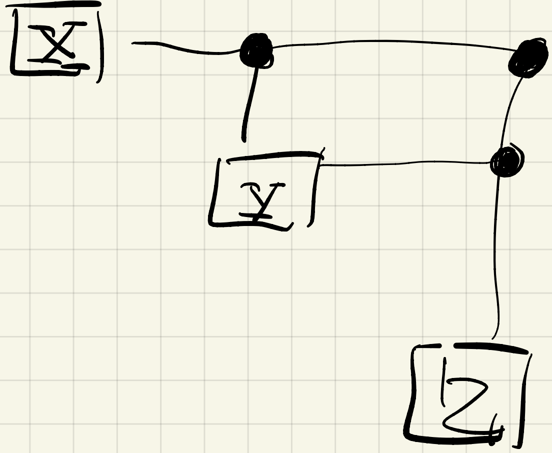
equivalance of NLBGS

to chain rule

$x = X()$

$y = Y(x)$

$z = Z(x, y)$



$$\overset{(n,e)}{\left[\frac{dz}{dx}\right]} = \overset{(n,e)}{\left[\frac{\partial z}{\partial x}\right]} + \underbrace{\overset{(n,m)}{\left[\frac{\partial z}{\partial y}\right]}\overset{(m,e)}{\left[\frac{dy}{dx}\right]}}_{(n,e)}$$

$x \Rightarrow$ size e

$y \Rightarrow$ size m

$z \Rightarrow$ size n

Basic Chain rule

Transform eqn. to use UDE

$$U = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad R(U) = r = \begin{bmatrix} x - x^* \\ y - Y(x) \\ z - Z_1(x,y) \end{bmatrix}$$

unknown vector

residual form

UDE linear system, matrix eqn.

$$\left[\frac{\partial R}{\partial U}\right]\left[\frac{dU}{dx}\right] = \begin{bmatrix} [I] \\ 0 \\ 0 \end{bmatrix}$$

(e.g.)

solve for total derivs one column of RHS at a time

$$\frac{\partial R_x}{\partial x} = [I] \qquad \frac{\partial R_x}{\partial y} = \frac{\partial R_x}{\partial z} = 0$$

$$\frac{\partial R_y}{\partial x} = -\left[\frac{\partial Y}{\partial x}\right] \qquad \frac{\partial R_y}{\partial y} = [I] \qquad \frac{\partial R_y}{\partial z} = 0$$

$$\frac{\partial R_z}{\partial x} = -\left[\frac{\partial Z_1}{\partial x}\right] \qquad \frac{\partial R_z}{\partial y} = -\left[\frac{\partial Z_1}{\partial y}\right] \qquad \frac{\partial R_z}{\partial z} = [I]$$

$$\begin{array}{cc} & \begin{array}{ccc} x & \quad y & \quad z \end{array} \\ \begin{array}{c} x \\ \\ y \\ \\ z \end{array} & \begin{bmatrix} I & 0 & 0 \\ -\frac{\partial y}{\partial x} & I & 0 \\ -\frac{\partial z}{\partial x} & -\frac{\partial z}{\partial y} & I \end{bmatrix} \end{array} \begin{bmatrix} \frac{dx}{dx} \\ \frac{dy}{dx} \\ \frac{dz}{dx} \end{bmatrix} = \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\frac{\partial R}{\partial U} \qquad \frac{dU}{dx} \qquad \text{rhs}$$

Block GS algorithm:

while $\text{norm}\left(\left[\frac{\partial R}{\partial U}\right]\left[\frac{dU}{dx}\right] - \begin{bmatrix} I \\ 0 \\ \vdots \end{bmatrix}\right) > \text{tol}$:

for j in $0 \ldots$ n_comps:

1) for i in $0 \ldots$ n_inputs: only wrt inputs

$$\text{rhs}_j = \text{rhs}_j - \left[\frac{\partial R_j}{\partial U_i}\right]\left[\frac{dU_i}{dX}\right]$$

2) solve for $\left[\frac{dU_j}{dx}\right]$ via:

only wrt outputs

$$\left[\frac{\partial R_j}{\partial U_j}\right]\left[\frac{dU_j}{dx}\right] = \text{rhs}_j$$

$$
\begin{array}{c|c}
 & \begin{array}{ccc} x & y & z \end{array} \\
\begin{array}{c} x \\ y \\ z \end{array} &
\begin{bmatrix} I & 0 & 0 \\ -\frac{\partial Y}{\partial x} & I & 0 \\ -\frac{\partial Z_1}{\partial x} & -\frac{\partial Z_1}{\partial y} & I \end{bmatrix}
\end{array}
\begin{bmatrix} \frac{dx}{dx} \\ \frac{dy}{dx} \\ \frac{dz}{dx} \end{bmatrix}
=
\begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix}
$$

$$\frac{\partial R}{\partial U} \qquad \frac{dU}{dx} \qquad \text{rhs}$$

note: each block represents one comp!

Block GS algorithm:

While norm$\left(\left[\frac{\partial R}{\partial U}\right]\left[\frac{dU}{dx}\right] - \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right) >$ tol:

for j in 0... n_comps:

1) for i in 0... n_inputs:

only wrt inputs

In om: apply_linear

$$\text{rhs}_j = \text{rhs}_j - \left[\frac{\partial R_j}{\partial U_i}\right]\left[\frac{dU_i}{dx}\right]$$

2) solve for $\left[\frac{dU_j}{dx}\right]$ via:

In om: Solve_linear

only wrt outputs

$$\left[\frac{\partial R_j}{\partial U_j}\right]\left[\frac{dU_j}{dx}\right] = \text{rhs}_j$$

$$
\begin{array}{cccc}
 & x & y & z \\
x & I & 0 & 0 \\
y & -\frac{\partial Y}{\partial x} & I & 0 \\
z & -\frac{\partial Z}{\partial x} & -\frac{\partial Z}{\partial y} & I
\end{array}
\quad
\begin{bmatrix} \frac{dx}{dx} \\ \frac{dy}{dx} \\ \frac{dz}{dx} \end{bmatrix}
=
\begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix}
$$

$\frac{\partial R}{\partial U}$ $\quad$ $\frac{dU}{dx}$ $\quad$ rhs

Block GS algorithm:

note: each block represents one comp!

for j in 0... n_comps:

1) for i in 0... n_inputs: only wrt inputs

$$rhs_j = rhs_j - \left[\frac{\partial R_j}{\partial U_i}\right]\left[\frac{dU_i}{dX}\right]$$

2) solve for $\left[\frac{dU_j}{dX}\right]$ via:

only wrt outputs

$$\left[\frac{\partial R_j}{\partial U_j}\right]\left[\frac{dU_j}{dx}\right] = rhs_j$$

j=0 (X comp)

1) $\frac{\partial R_x}{\partial y} = 0 \quad \frac{\partial R_x}{\partial z} = 0 \Rightarrow [I] = [I] - 0$

2) $\left[\frac{\partial R_x}{\partial x}\right]\left[\frac{dx}{dx}\right] = [I] \Rightarrow \left[\frac{dx}{dx}\right] = [I]$ duh

j=1 (Y comp)

1) $\frac{\partial R_y}{\partial x} = -\frac{\partial Y}{\partial x}$; $\frac{\partial R_y}{\partial z} = 0 \Rightarrow rhs_y = 0 - \left[-\frac{\partial Y}{\partial x}\right]\left[\cancel{\frac{dx}{dx}}\right] \to [I]$

2) $\left[\frac{\partial R_y}{\partial y}\right]\left[\frac{dy}{dx}\right] = rhsy \Rightarrow [I]\left[\frac{dy}{dx}\right] = \left[\frac{\partial Y}{dx}\right][I]$

Now look at y row of ODE

$$\left[-\frac{\partial Y}{\partial X}\right]\left[\frac{dX}{dX}\right] + [I]\left[\frac{dy}{dx}\right] = 0$$

$$[I]\left[\frac{dy}{dx}\right] = \left[\frac{\partial Y}{dx}\right][I]$$

same eqn!

$$
\begin{array}{c|ccc}
 & x & y & z \\
\hline
x & I & 0 & 0 \\
y & -\frac{\partial y}{\partial x} & I & 0 \\
z & -\frac{\partial z_1}{\partial x} & -\frac{\partial z_1}{\partial y} & I
\end{array}
\quad
\begin{bmatrix} \frac{dx}{dx} \\ \frac{dy}{dx} \\ \frac{dz}{dx} \end{bmatrix}
=
\begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix}
$$

$$\underbrace{\frac{\partial R}{\partial U}}\qquad\qquad \underbrace{\frac{dU}{dx}} \qquad\qquad \text{rhs}$$

note: each block represents one comp!

Block GS algorithm:

for j in $0 \dots$ n_comps:

1) for i in $0 \dots$ n_inputs: only wrt inputs

$$\text{rhs}_j = \text{rhs}_j - \left[\frac{\partial R_j}{\partial U_i}\right]\left[\frac{dU_i}{dx}\right]$$

2) solve for $\left[\frac{dU_j}{dx}\right]$ via:

only wrt outputs

$$\left[\frac{\partial R_j}{\partial U_j}\right]\left[\frac{dU_j}{dx}\right] = \text{rhs}_j$$

$j=2$ (z_1 comp)

1) $\frac{\partial R_z}{\partial x} = -\frac{\partial z_1}{\partial x}$ $\quad \frac{\partial R_z}{\partial y} = -\frac{\partial z_1}{\partial y}$ $\Rightarrow \text{rhs}_z = 0 - -\left[\frac{\partial z_1}{\partial x}\right]\cancel{\frac{dx}{dx}}\,[I] - -\left[\frac{\partial z_1}{\partial y}\right]\left[\frac{dy}{dx}\right]$

$$\text{rhs}_z = \left[\frac{\partial z_1}{\partial x}\right] + \left[\frac{\partial z_1}{\partial y}\right]\left[\frac{dy}{dx}\right]$$

2) $$\cancel{\left[\frac{\partial R_{z_1}}{\partial z}\right]} \to [I] \quad \left[\frac{dz}{dx}\right] = \left[\frac{\partial z_1}{\partial x}\right] + \left[\frac{\partial z_1}{\partial y}\right]\left[\frac{dy}{dx}\right]$$

$$\left[\frac{dz}{dx}\right] = \left[\frac{\partial z_1}{\partial x}\right] + \left[\frac{\partial z_1}{\partial y}\right]\left[\frac{dy}{dx}\right]$$

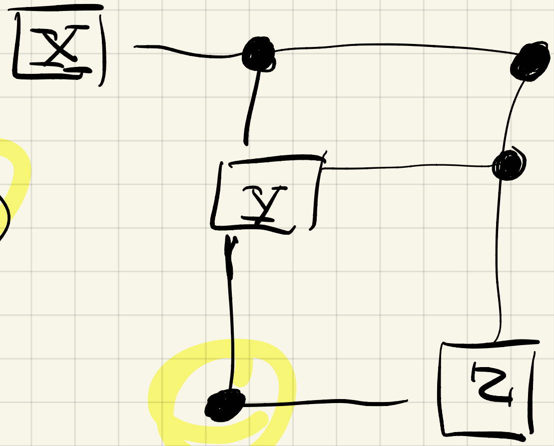
same equation as regular chain rule

What if you add some coupling?

$x = X()$

$y = Y(x, z)$

$z = Z(x, y)$



Now how do you chain rule this?

its possible, but a little tricky...

for another lecture

UDE form requires only a minor adjustment

$R_x = x - x^*$

$R_y = y - Y(x, z)$

$R_z = z - Z(x, y)$

$$\left[\frac{\partial R}{\partial U}\right]\left[\frac{dU}{dx}\right] = \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$\Longleftarrow$ same

$$\begin{bmatrix} I & 0 & 0 \\ -\frac{\partial Y}{\partial x} & I & -\frac{\partial Y}{\partial z} \\ \frac{\partial Z}{\partial x} & \frac{\partial Z}{\partial y} & I \end{bmatrix} \begin{bmatrix} \quad \end{bmatrix} = \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$