HW4

Monday, February 17, 2025 1:18 PM

4.1 Derivatives using Complex Step

Calculate the Jacobian for the 4-dimensional Rosenbrock function (see Equation D.3) using the complex step at the point x = [0.5, -0.5, 0.5, 1], with h going from 0.1, 0.001, 0.001, ... 1e-30. This means there are 4 design variables, and therefore the gradient will be 4x1 and the Jacobian would be 1x4. Compare this to the answer with forward finite differencing using the same values of h = 0.1, 0.001, 0.001, ... 1e-30. Compare both the complex and forward step to the exact derivative in a plot similar to Fig 6.9 in the textbook. (Note, you will have to decide how to present the error between the Jacobians).

The Rosenbrock function can be extended to n dimensions by defining the sum,

$$f(x) = \sum_{i=1}^{n-1} \left(100\left(x_{i+1} - x_i{}^2\right)^2 + (1 - x_i)^2\right). \qquad \text{(D.3)}$$

579

4D Rosenbrock

f(x1,x2,x3,x4) = 100*(x2-x1**2)**2 + (1-x1)**2 + 100*(x3-x2**2)**2 + (1-x2)**2 + 100*(x4 - x3**2)**2 + (1-x3)**2

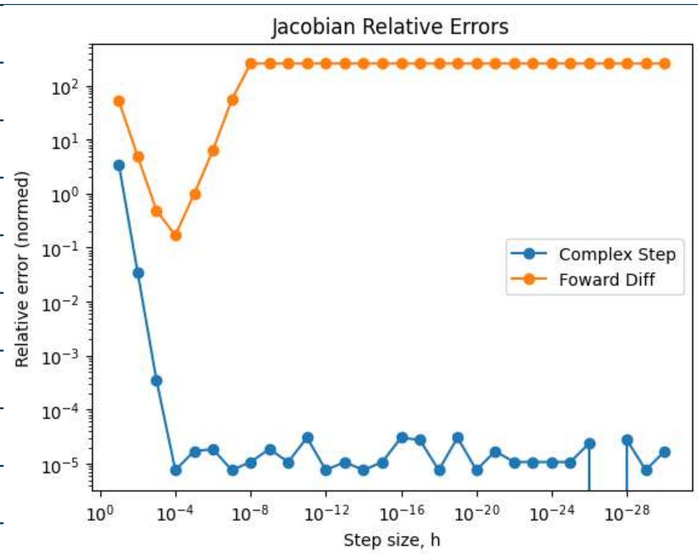
Dfdx = 4x1

Jacobian = 1x4

Example $f(x) = \begin{bmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{bmatrix} = \begin{bmatrix} x_1 x_2 + \sin x_1 \\ x_1 x_2 + x_2^2 \end{bmatrix}$

$$J = \begin{array}{cc} & x \longrightarrow \\ f \downarrow & \begin{bmatrix} x_2 + \cos x_1 & x_1 \\ x_2 & x_1 + 2x_2 \end{bmatrix} \end{array}$$

complex step $\dfrac{\partial f}{\partial x_j} = \dfrac{\mathrm{Im}\left(f(x + ih\hat{e}_j)\right)}{h} + O(h^2)$



4.2 Derivatives using AD

Solve for the Jacobian at the point $(x = 2, y = 1.5)$, for the vector function $f = [f_1, f_2]^T$ presented in the following python code. Show your work as needed (by hand or with code).

```
3  a = x**2 + 3*y**2
4  b = sin(a)
5  c = 3*x**2 + y**2
6  d = sin(c)
7  e = x**2 + y**2
8  g = -0.5*exp(-e/2)
9  f1 = a + c + g
10 h = x*y
11 f2 = a - g + h
```

$$\dot{v}_i = \sum_{k=j}^{i-1} \frac{\partial v_i}{\partial v_k} \dot{v}_k$$

$$J = \begin{array}{c} \\ f_1 \\ f_2 \end{array} \begin{array}{cc} x & y \\ \multicolumn{2}{c}{\left[\begin{array}{cc} \quad & \quad \\ \quad & \quad \end{array}\right]} \end{array}$$

@ $j = k = 1$ ie x

	i	$\dot{v}_i$
$v_1 = x$	1	$\dot{v}_1 = \frac{\partial v_1}{\partial v_1} \dot{v}_1 = 1$
$v_2 = y$	2	$\dot{v}_2 = \cancel{\frac{\partial v_2}{\partial v_1}}^{0} \dot{v}_1 = 0$
$v_3 = v_1^2 + 3v_2^2$	3	$\dot{v}_3 = \frac{\partial v_3}{\partial v_2} \cancel{\dot{v}_2}^{0} + \frac{\partial v_3}{\partial v_1} \dot{v}_1 = 2v_1$
$v_4 = \sin(v_3)$	4	$\dot{v}_4 = \frac{\partial v_4}{\partial v_3} \dot{v}_3 = \cos(v_3)(2v_1)$
$v_5 = 3v_1^2 + v_2^2$	5	$\dot{v}_5 = \frac{\partial v_5}{\partial v_2} \cancel{\dot{v}_2}^{0} + \frac{\partial v_5}{\partial v_1} \dot{v}_1 = 6v_1(1)$
$v_6 = \sin(v_5)$	6	$\dot{v}_6 = \frac{\partial v_6}{\partial v_5} \dot{v}_5 = \cos(v_5)(6v_1)$
$v_7 = v_1^2 + v_2^2$	7	$\dot{v}_7 = \frac{\partial v_7}{\partial v_2} \cancel{\dot{v}_2}^{0} + \frac{\partial v_7}{\partial v_1} \dot{v}_1 = 2v_1(1)$
$\ldots e^{(-v_7/2)}$		$\ldots e^{(-v_7/2)} \ldots$

$v_7 = v_1^2 + v_2^2$	7	$\dot{v}_7 = \frac{\partial v_7}{\partial v_2}\dot{v}_2 + \frac{\partial v_7}{\partial v_1}\dot{v}_1 = 2v_1\,(1)$
$v_8 = -0.5e^{\wedge(-v_7/2)}$	8	$\dot{v}_8 = \frac{\partial v_8}{\partial v_7}\dot{v}_7 = \frac{1}{4}e^{\wedge(-v_7/2)}\,(2v_1)$
$v_9 = v_3 + v_5 + v_8 = f_1$	9	$\dot{v}_9 = \frac{\partial v_9}{\partial v_8}\dot{v}_8 + \frac{\partial v_9}{\partial v_5}\dot{v}_5 + \frac{\partial v_9}{\partial v_3}\dot{v}_3 = v_1/2\,e^{\wedge(-v_7/2)} + 8v_1 = \frac{df_1}{dx_1}$
$v_{10} = v_1 v_2$	10	$\dot{v}_{10} = \frac{\partial v_{10}}{\partial v_2}\cancel{\dot{v}_2}^{0} + \frac{\partial v_{10}}{\partial v_1}\dot{v}_1 = v_2$
$v_{11} = v_3 - v_8 + v_{10} = f_2$	11	$\dot{v}_{11} = \frac{\partial v_{11}}{\partial v_{10}}\dot{v}_{10} + \frac{\partial v_{11}}{\partial v_8}\dot{v}_8 + \frac{\partial v_{11}}{\partial v_3}\dot{v}_3 = v_2 - \frac{v_1}{2}e^{\wedge(-v_7/2)} + 2v_1$

df_2/dx_1

	i	$\dot{v}_i$ @ $j=k=2$ ie y
$v_1 = x$	1	$\dot{v}_1 = 0$
$v_2 = y$	2	$\dot{v}_2 = 1$
$v_3 = v_1^2 + 3v_2^2$	3	$\dot{v}_3 = \frac{\partial v_3}{\partial v_2}\dot{v}_2 + \frac{\partial v_3}{\partial v_1}\cancel{\dot{v}_1}^{0} = 6v_2\,(1)$
$v_4 = \sin(v_3)$	4	$\dot{v}_4 = \frac{\partial v_4}{\partial v_3}\dot{v}_3 = \cos(v_3)\,6v_2$
$v_5 = 3v_1^2 + v_2^2$	5	$\dot{v}_5 = \frac{\partial v_5}{\partial v_2}\dot{v}_2 + \frac{\partial v_5}{\partial v_1}\cancel{\dot{v}_1}^{0} = 2v_2\,(1)$
$v_6 = \sin(v_5)$	6	$\dot{v}_6 = \frac{\partial v_6}{\partial v_5}\dot{v}_5 = \cos(v_5)\,2v_2$
$v_7 = v_1^2 + v_2^2$	7	$\dot{v}_7 = \frac{\partial v_7}{\partial v_2}\dot{v}_2 + \frac{\partial v_7}{\partial v_1}\cancel{\dot{v}_1}^{0} = 2v_2\,(1)$
$v_8 = -0.5e^{\wedge(-v_7/2)}$	8	$\dot{v}_8 = \frac{\partial v_8}{\partial v_7}\dot{v}_7 = \frac{1}{4}e^{(-v_7/2)}\,2v_2 = \frac{v_2}{2}e^{\wedge(-v_7/2)}$
$v_9 = v_3 + v_5 + v_8 = f_1$	9	$\dot{v}_9 = \frac{\partial v_9}{\partial v_8}\dot{v}_8 + \frac{\partial v_9}{\partial v_5}\dot{v}_5 + \frac{\partial v_9}{\partial v_3}\dot{v}_3 = \frac{v_2}{2}e^{\wedge(-v_7/2)} + 8v_2$ df_1/dx_2
$v_{10} = v_1 v_2$	10	$\dot{v}_{10} = \frac{\partial v_{10}}{\partial v_2}\dot{v}_2 + \frac{\partial v_{10}}{\partial v_1}\cancel{\dot{v}_1}^{0} = v_1\,(1)$
$v_{11} = v_3 - v_8 + v_{10} = f_2$	11	$\dot{v}_{11} = \frac{\partial v_{11}}{\partial v_{10}}\dot{v}_{10} + \frac{\partial v_{11}}{\partial v_8}\dot{v}_8 + \frac{\partial v_{11}}{\partial v_3}\dot{v}_3 = v_1 - \frac{v_2}{2}e^{\wedge(-v_7/2)} + 6v_2$

$$v_{11} = v_3 - v_8 + v_{10} = f_2 \quad || \quad \dot{v}_{11} = \frac{\partial v_{11}}{\partial v_{10}} \dot{v}_{10} + \frac{\partial v_{11}}{\partial v_8} \dot{v}_8 + \frac{\partial v_{11}}{\partial v_3} \dot{v}_3 = \underbrace{\left(v_1 - \frac{v_2}{2} e \quad + 6 v_2 \right)}_{df_2/dx_2 =}$$

$$J = \begin{array}{c} \\ f_1 \\ f_2 \end{array} \begin{array}{c} \begin{array}{cc} x & y \end{array} \\ \begin{bmatrix} \frac{v_1}{2} \exp(-v_7/2) + 8v_1 & \frac{v_2}{2} \exp(-v_7/2) + 8v_2 \\ v_2 - v_1/2 \exp(-v_7/2) + 2v_1 & v_1 - \frac{v_2}{2} \exp(-v_7/2) + 6v_2 \end{bmatrix} \end{array}$$

$$\boxed{J = \begin{array}{c} \\ f_1 \\ f_2 \end{array} \begin{array}{c} \begin{array}{cc} x & y \end{array} \\ \begin{bmatrix} 16.04 & 12.03 \\ 5.45 & 10.97 \end{bmatrix} \end{array}}$$

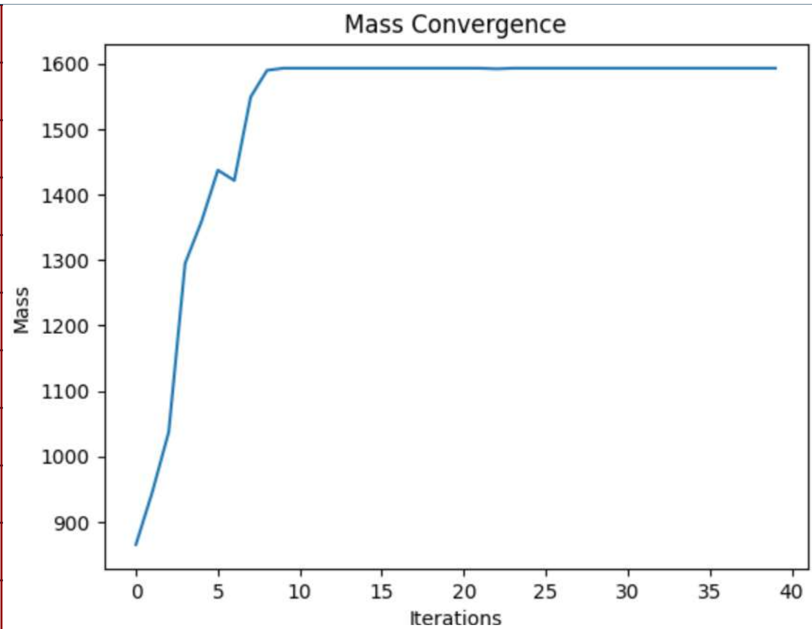
← different order than jax.jacobian?

4.3 Jacobian of mass with respect to area, and of stress (sigma) with respect to area:

```
Finite Difference Relative Error (dmdA):              0.0004752932
Complex Step Relative Error (dmdA):                   5.6640044e-08
AD Relative Error (dmdA):                             0.0
Finite Difference Stress Relative Error (dsigmadA):  0.2149642
Complex Step Stress Relative Error (dsigmadA):        0.501
AD Stress Relative Error (dsigmadA):                  6.67646e-07
```

Discussion: For the dmdA Jacobians, forward difference performed the worse, but still better than expected. Complex did very well and AD is the best we could do. The stress Jacobians were more difficult, and I expected finite difference to be the furthest off. This ended up not being the case, and complex step was further off. It is possible that I implemented something wrong with Complex step in this case, especially as it was difficult to complexify everything. AD did very well as expected.
Merits: Finite difference is the easiest to implement, but was quite touchy depending on the step size, and for stress didn't do well. Complex step does quite well but is hard to implement which can lead to errors.

4.4 Truss Optimization With the provided gradients you can now optimize the truss. The problem is identical to Homework #1, except this time ***you will supply the derivatives***. Solve this optimization problem using **one** of the exact derivative methods (not finite differencing). Generate a convergence plot and report the number of function calls to the truss function required to converge. Discuss your findings in 200 words or less.



Function Calls: 362

Discussion: It is cool that you can provide your own gradients for everything and in theory it should be faster. I think I had a bug somewhere, so mine, in the end, had more function calls (362 vs 298) but not a ton, although I did converge to the wrong value as well. All in all, I think my take away is that it could be helpful, but you would need to be very careful.