Exercise sheet: Auto-diff

Let $\mathbf{F}$ be a vector-valued function that maps from $\mathbb{R}^3$ to $\mathbb{R}^2$,

$$
\begin{aligned}
y_1 &= f_1(x_1, x_2, x_3) = x_1 x_3 + \log(x_2 + x_1) \times e^{-x_3} \\
y_2 &= f_2(x_1, x_2, x_3) = e^{-x_2} + \cos(x_1 x_3).
\end{aligned}
$$

1. (*) Compute the Jacobian using manual differentiation and evaluate the Jacobian at the point $(x_1 = 3, x_2 = 5, x_3 = 1)$

   Answer The Jacobian $\mathbf{J}$ has dimensions 2×3 and is given as

$$
\mathbf{J} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \end{bmatrix},
$$

   where the derivatives in the first row are given as

$$
\begin{aligned}
\frac{df_1}{dx_1} &= x_3 + e^{-x_3} \frac{1}{(x_2 + x_1)}, \\
\frac{df_1}{dx_2} &= e^{-x_3} \frac{1}{x_2 + x_1}, \\
\frac{df_1}{dx_3} &= x_1 - \log(x_2 + x_1) e^{-x_3},
\end{aligned}
$$

   and the derivatives in the second row are given as

$$
\begin{aligned}
\frac{df_2}{dx_1} &= -x_3 \sin(x_1 x_3) \\
\frac{df_2}{dx_2} &= -e^{-x_2} \\
\frac{df_2}{dx_3} &= -x_1 \sin(x_1 x_3).
\end{aligned}
$$

   The Jacobian at the point $(x_1 = 3, x_2 = 5, x_3 = 1)$ is given as

$$
\mathbf{J}(x_1 = 3, x_2 = 5, x_3 = 1) = \begin{bmatrix} 1.0459849301 & 0.0459849301 & 2.2350162077 \\ -0.1411200081 & -0.006737947 & -0.4233600242 \end{bmatrix}.
$$

2. (*) Compute the Jacobian at the same point that in the previous point, but using finite difference approximation.

   Answer We can compute the finite difference approximations for the partial derivatives using the Python code shown in following snippet:

```python
import numpy as np

def f1(x1, x2, x3):
  return x1*x3 + np.log(x2+x1)*np.exp(-x3)

def f2(x1, x2, x3):
  return np.exp(-x2) + np.cos(x1*x3)

x1_0 = 3
x2_0 = 5
x3_0 = 1
epsilon = 1e-6

df1dx1_numerical = (f1(x1_0+epsilon, x2_0, x3_0) - f1(x1_0, x2_0, x3_0))/epsilon
df1dx2_numerical = (f1(x1_0, x2_0+epsilon, x3_0) - f1(x1_0, x2_0, x3_0))/epsilon
df1dx3_numerical = (f1(x1_0, x2_0, x3_0+epsilon) - f1(x1_0, x2_0, x3_0))/epsilon

df2dx1_numerical = (f2(x1_0+epsilon, x2_0, x3_0) - f2(x1_0, x2_0, x3_0))/epsilon
df2dx2_numerical = (f2(x1_0, x2_0+epsilon, x3_0) - f2(x1_0, x2_0, x3_0))/epsilon
df2dx3_numerical = (f2(x1_0, x2_0, x3_0+epsilon) - f2(x1_0, x2_0, x3_0))/epsilon
```

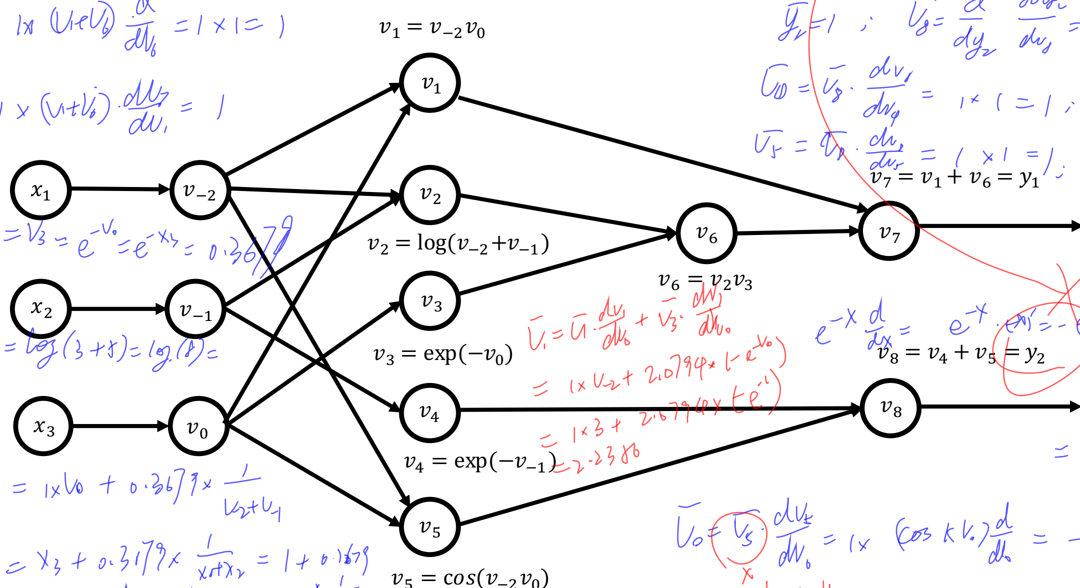
The Jacobian computed using finite differences is approximated as

$$\mathbf{J}(x_1 = 3, x_2 = 5, x_3 = 1) \approx \begin{bmatrix} 1.0459849276 & 0.0459849274 & 2.2350165896 \\ -0.1411195131 & -0.0067379436 & -0.4233555692 \end{bmatrix}.$$

$\dot{x}_1 = 1.0460$
$\dot{x}_2 = 0.0392$
$\dot{x}_3 = -0.4234$

Should only calculate one line at a time (for Jacobian) y_2:

3. (**) Draw the computational graph.



$\bar{v}_1 = \bar{v}_7 \cdot \frac{dv_7}{dv_1} = 1 \times v_6$

$\bar{y}_2 = 1$; $\bar{v}_8 = \frac{d}{dy_2}\frac{dy_2}{dv_8} = \bar{y}_2 \cdot x \; 1 = 1$

$\bar{v}_6 = \bar{v}_7 \cdot \frac{dv_7}{dv_6} = 1 \times (v_1+v_6)\frac{d}{dv_6} = 1 \times 1 = 1$

$\bar{v}_1 = \bar{v}_7 \cdot \frac{dv_7}{dv_1} = 1 \times (v_1+v_6)\cdot\frac{dv_7}{dv_1} = 1$

$\bar{v}_2 = \bar{v}_6 \cdot \frac{dv_6}{dv_2} = 1 \times v_3 = \bar{v}_3 = e^{-v_0} = e^{-x_3} = 0.3679$

$\bar{v}_3 = \bar{v}_6 \cdot \frac{dv_6}{dv_3} = 1 \times v_2 = v_2 = \log(3+5) = \log(8) = 2.0794$

$\bar{v}_2 = \bar{v}_1 \cdot \frac{dv_1}{dv_2} + \bar{v}_2 \cdot \frac{dv_2}{dv_2} = 1 \times v_0 + 0.3679 \times \frac{1}{v_2+v_1}$

$= 1.0460$

$= x_3 + 0.3179 \times \frac{1}{x_1+x_2} = 1 + 0.1679 \times \frac{1}{8}$

$\frac{d}{dx_1} = \frac{d}{dv_2}\cdot\frac{dv_2}{dx_1} = \frac{d}{dv_2}\cdot x_1 \cdot dx_1 = \frac{d}{dv_2} = 1.0460$

$\bar{v}_0 = \bar{v}_8 \cdot \frac{dv_8}{dv_4} = 1 \times 1 = 1$; $\bar{v}_4 = \bar{v}_2 \cdot \frac{dv_2}{dv_4} + \bar{v}_3 \cdot \frac{dv_3}{dv_4}$

$\bar{v}_5 = \bar{v}_8 \cdot \frac{dv_8}{dv_5} = 1 \times 1 = 1$;

$= 0.3679 \times \frac{1}{v_2+v_4} + 1 \times -e^{-v_4}$

$= 0.3679 \times \frac{1}{8} - e^{-5} = 0.0312$

$\bar{v}_1 = \bar{v}_1 \cdot \frac{dv_1}{dv_0} + \bar{v}_3 \cdot \frac{dv_3}{dv_0}$

$= 1 \times v_2 + 2.0794 \times (-e^{-v_0})$

$= 1 \times 3 + 2.0794 \times (-e^{-1})$

$= 2.2350$

$e^{-x}\frac{d}{dx} = e^{-x}(x)' = -e^{-x}$

$\frac{df}{dx_2} = \frac{df}{dv_4}\cdot\frac{dv_4}{dx_2}$

$= \bar{v}_4 \cdot x_2 \cdot \frac{d}{dx_2} = \bar{v}_4 = 0.0312$

$\bar{v}_0 = \bar{v}_5 \cdot \frac{dv_5}{dv_0} = 1 \times (\cos k v_0)\frac{d}{dv_0} = -k\sin k v_0 = -v_2 \sin v_2 v_0$

for y_1

$= -0.4234$

Figure 1: Computational graph for the vector-valued function

$\frac{df}{dx_3} = \frac{df}{dv_0}\cdot\frac{dv_0}{dx_3} = \bar{v}_0 \cdot x_3\frac{d}{dx_3} = \bar{v}_0$

4. (***) Compute the Jacobian using AD in forward mode. Write the expressions for all the intermediate variables $\dot{v}_i$ in the forward tangent trace.

Forward primal trace	Forward tangent trace
$v_{-2} = x_1$	$\dot{v}_{-1} = \dot{x}_1$
$v_{-1} = x_2$	$\dot{v}_{-1} = \dot{x}_2$
$v_0 \; = x_3$	$\dot{v}_0 \; = \dot{x}_3$
$v_1 \; = v_{-2}v_0$	$\dot{v}_1 \; = v_0$
$v_2 \; = \log(v_{-2} + v_{-1})$	$\dot{v}_2 \; = 1/(v_{-2} + v_{-1})$
$v_3 \; = \exp(-v_0)$	$\dot{v}_3 \; = 0$
$v_4 \; = \exp(-v_{-1})$	$\dot{v}_4 \; = 0$
$v_5 \; = \cos(v_{-2}v_0)$	$\dot{v}_5 \; = -v_0 \sin(v_{-2}v_0)$
$v_6 \; = v_2 v_3$	$\dot{v}_6 \; = v_3 \dot{v}_2$
$v_7 \; = v_1 + v_6$	$\dot{v}_7 \; = \dot{v}_1 + \dot{v}_6$
$v_8 \; = v_4 + v_5$	$\dot{v}_8 \; = \dot{v}_5$
$y_1 \; = v_7$	$\dot{y}_1 \; = \dot{v}_7$
$y_2 \; = v_8$	$\dot{y}_2 \; = \dot{v}_8$

Table 1: Forward tanget trace for $\frac{dy_1}{dx_1}$ and $\frac{dy_2}{dx_1}$

Forward primal trace		Forward tangent trace	
$v_{-2} = x_1$	$= 3$	$\dot{v}_{-1} = \dot{x}_1$	$= 1$
$v_{-1} = x_2$	$= 5$	$\dot{v}_{-1} = \dot{x}_2$	$= 0$
$v_0 \; = x_3$	$= 1$	$\dot{v}_0 \; = \dot{x}_3$	$= 0$
$v_1 \; = v_{-2}v_0$	$= 3$	$\dot{v}_1 \; = v_0$	$= 1$
$v_2 \; = \log(v_{-2} + v_{-1})$	$= 2.079$	$\dot{v}_2 \; = 1/(v_{-2} + v_{-1})$	$= 0.125$
$v_3 \; = \exp(-v_0)$	$= 0.367$	$\dot{v}_3 \; = 0$	
$v_4 \; = \exp(-v_{-1})$	$= 0.006$	$\dot{v}_4 \; = 0$	
$v_5 \; = \cos(v_{-2}v_0)$	$= -0.989$	$\dot{v}_5 \; = -v_0 \sin(v_{-2}v_0)$	$= -0.141$
$v_6 \; = v_2 v_3$	$= 0.764$	$\dot{v}_6 \; = v_3 \dot{v}_2$	$= 0.045$
$v_7 \; = v_1 + v_6$	$= 3.764$	$\dot{v}_7 \; = \dot{v}_1 + \dot{v}_6$	$= 1.045$
$v_8 \; = v_4 + v_5$	$= -0.983$	$\dot{v}_8 \; = \dot{v}_5$	$= -0.141$
$y_1 \; = v_7$	$= 3.764$	$\dot{y}_1 \; = \dot{v}_7$	$= 1.045$
$y_2 \; = v_8$	$= -0.983$	$\dot{y}_2 \; = \dot{v}_8$	$= -0.141$

$\bar{v}_6 = \bar{v}_7 \cdot \frac{dv_7}{dv_6} = 1 \times \frac{(v_1+v_6)d}{dv_6}$

$= 1$

Table 2: Derivatives for $\frac{dy_1}{dx_1}$ and $\frac{dy_2}{dx_1}$, at $(x_1 = 3, x_2 = 5, x_3 = 1)$

Answer Let us compute the forward tangent trace for $\frac{dy_1}{dx_1}$ and $\frac{dy_2}{dx_1}$. Table 1 shows the forward primal trace and the forward tangent trace.

We use table 1 to compute the derivatives $\frac{dy_1}{dx_1}$ and $\frac{dy_2}{dx_1}$ at $(x_1 = 3, x_2 = 5, x_3 = 1)$. Notice how table 2 provides the first column of the Jacobian.

The other two columns of the Jacobian can be computed using a similar procedure. This is left to the student to complete.

5. (***) Compute the Jacobian using AD in reverse mode. Write the expressions for all the adjoints $\bar{v}_i$ in the reverse derivative trace.

Answer. Let us compute the partial derivatives $\frac{\partial y_1}{\partial x_1}$, $\frac{\partial y_1}{\partial x_2}$ and $\frac{\partial y_1}{\partial x_3}$ using the reverse mode. The computation of $\frac{\partial y_2}{\partial x_1}$, $\frac{\partial y_2}{\partial x_2}$ and $\frac{\partial y_2}{\partial x_3}$ is left to the student.

The adjoint $\bar{y}_1$ is simply $\bar{y}_1 = 1$.

Looking at the computational graph, we now compute $\bar{v}_7 = \frac{\partial y_1}{\partial v_7} = 1$.

The adjoints we need to compute are then

$$\begin{aligned}
\bar{v}_6 &= \bar{v}_7 \frac{\partial v_7}{\partial v_6} = \bar{v}_7(1) = 1 \\
\bar{v}_1 &= \bar{v}_7 \frac{\partial v_7}{\partial v_1} = \bar{v}_7(1) = 1 \\
\bar{v}_2 &= \bar{v}_6 \frac{\partial v_6}{\partial v_2} = \bar{v}_6 v_3 = (1)(0.367) = 0.367 \\
\bar{v}_3 &= \bar{v}_6 \frac{\partial v_6}{\partial v_3} = \bar{v}_6 v_2 = (1)(2.079) = 2.079 \\
\bar{v}_{-2} &= \bar{v}_1 \frac{v_1}{v_{-2}} + \bar{v}_2 \frac{v_2}{v_{-2}} = \bar{v}_1 v_0 + \bar{v}_2 \frac{1}{v_{-2} + v_{-1}} = (1)(1) + (0.367)/(8) = 1.0458 \\
\bar{v}_{-1} &= \bar{v}_2 \frac{v_2}{v_{-1}} = \bar{v}_2 \frac{1}{v_{-2} + v_{-1}} = (0.367)/(8) = 0.0459 \\
\bar{v}_0 &= \bar{v}_1 \frac{v_1}{v_0} + \bar{v}_3 \frac{v_3}{v_0} = \bar{v}_1 v_{-2} + \bar{v}_3(-\exp(-v_0)) = (1)(3) - (2.079)\exp(-1) = 2.235.
\end{aligned}$$

Finally, we get

$$\begin{aligned}
\bar{x}_1 &= \bar{v}_{-2} \frac{\partial v_{-2}}{\partial x_1} = \bar{v}_{-2} = 1.0458 \\
\bar{x}_2 &= \bar{v}_{-1} \frac{\partial v_{-1}}{\partial x_2} = \bar{v}_{-1} = 0.0459 \\
\bar{x}_3 &= \bar{v}_0 \frac{\partial v_0}{\partial x_3} = \bar{v}_0 = 0.0459
\end{aligned}$$