Reverse Mode AD

To Understand the Reverse Mode, We Again Turn to Our Simple Function

$$\begin{aligned} f(x_1, x_2) &= x_1^2 + x_2 \sin\left(x_1^2\right) \\ &= v_1 + v_2 \end{aligned}$$

The total derivative of f with respect to x_i $(i = 1, 2)$ is

$$\frac{df}{dx_i} = \frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial v_1}\frac{\partial v_1}{\partial x_i} + \frac{\partial f}{\partial v_2}\left(\frac{\partial v_2}{\partial v_1}\frac{\partial v_1}{\partial x_i} + \frac{\partial v_2}{\partial x_i}\right)$$

The Forward Mode Evaluates Partial Derivatives Right-to-Left

Step 1: $$\frac{df}{dx_i} = \frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial v_1}\frac{\partial v_1}{\partial x_i} + \frac{\partial f}{\partial v_2}\left(\frac{\partial v_2}{\partial v_1}\frac{\partial v_1}{\partial x_i} + \frac{\partial v_2}{\partial x_i}\right)$$

Step 2: $$\frac{df}{dx_i} = \frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial v_1}\frac{\partial v_1}{\partial x_i} + \frac{\partial f}{\partial v_2}\left(\frac{\partial v_2}{\partial v_1}\frac{\partial v_1}{\partial x_i} + \frac{\partial v_2}{\partial x_i}\right)$$

Step 3: $$\frac{df}{dx_i} = \frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial v_1}\frac{\partial v_1}{\partial x_i} + \frac{\partial f}{\partial v_2}\left(\frac{\partial v_2}{\partial v_1}\frac{\partial v_1}{\partial x_i} + \frac{\partial v_2}{\partial x_i}\right)$$

The Reverse Mode Evaluates Partial Derivatives Left-to-Right

Step 1: $$\frac{df}{dx_i} = \frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial v_1}\frac{\partial v_1}{\partial x_i} + \frac{\partial f}{\partial v_2}\left(\frac{\partial v_2}{\partial v_1}\frac{\partial v_1}{\partial x_i} + \frac{\partial v_2}{\partial x_i}\right)$$

Step 2: $$\frac{df}{dx_i} = \frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial v_1}\frac{\partial v_1}{\partial x_i} + \frac{\partial f}{\partial v_2}\left(\frac{\partial v_2}{\partial v_1}\frac{\partial v_1}{\partial x_i} + \frac{\partial v_2}{\partial x_i}\right)$$

Step 3: $$\frac{df}{dx_i} = \frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial v_1}\frac{\partial v_1}{\partial x_i} + \frac{\partial f}{\partial v_2}\left(\frac{\partial v_2}{\partial v_1}\frac{\partial v_1}{\partial x_i} + \frac{\partial v_2}{\partial x_i}\right)$$

To Make This Work, We Must Keep Track of the Total Derivatives

We will use an overbar on variables to denote the total derivative of f with respect to that variable

$$\bar{f} = \frac{df}{df} = 1$$

$$\bar{v}_1 = \frac{df}{dv_1} \qquad \bar{v}_2 = \frac{df}{dv_2}$$

$$\bar{x}_1 = \frac{df}{dx_1} \qquad \bar{x}_2 = \frac{df}{dx_2}$$

- $\bar{x}_1$ and $\bar{x}_2$ define the gradient we want

Let's Apply the Reverse Mode To Our Simple Function

Recall the Matlab implementation:

```
1 function [f] = func(x1, x2)
2 % compute a simple function value
3 v1 = x1.^2;
4 v2 = x2.*sin(v1);
5 f = v1 + v2;
6 end
```

- For the reverse mode, we first evaluate lines 3-5
- Then we step backwards through the code, differentiating line by line

Initialize $\bar{f} = 1$, $\bar{v}_1 = 0$, $\bar{v}_2 = 0$, $\bar{x}_1 = 0$, $\bar{x}_2 = 0$.

Let's Apply the Reverse Mode To Our Simple Function (cont.)

```
1 function [f] = func(x1, x2)
2 % compute a simple function value
3 v1 = x1.^2;
4 v2 = x2.*sin(v1);
5 f = v1 + v2;
6 end
```

Differentiating line 5 with respect to the variables on its right (i.e. v_1 and v_2) gives us

$$\bar{v}_2 =$$

$$\bar{v}_1 =$$

Let's Apply the Reverse Mode To Our Simple Function (cont.)

```
1 function [f] = func(x1, x2)
2 % compute a simple function value
3 v1 = x1.^2;
4 v2 = x2.*sin(v1);
5 f = v1 + v2;
6 end
```

Differentiating line 4 with respect to the variables on its right (x_2 and v_1) gives us

$$\bar{v}_1 =$$

$$\bar{x}_2 =$$

Let's Apply the Reverse Mode To Our Simple Function (cont.)

```
1 function [f] = func(x1, x2)
2 % compute a simple function value
3 v1 = x1.^2;
4 v2 = x2.*sin(v1);
5 f = v1 + v2;
6 end
```

Finally, differentiating line 3 with respect to the variables on its right (x_1) gives us

$$\bar{x}_1 =$$

Let's Apply the Reverse Mode To Our Simple Function (cont.)

Thus, in the end we have

$$\frac{df}{dx_1} = \bar{x}_1 = 2x_1(1 + x_2 \cos(v_1))$$
$$\frac{df}{dx_2} = \bar{x}_2 = \sin(v_1)$$

- You can easily verify that these are the correct values for the gradient of f evaluated at (x_1, x_2).

Putting This Into the Form of Matlab Code...

```
1  function [f, dfdx1, dfdx2] = dfunc_reverse(x1, x2)
2  % compute a simple function value and its gradient
3  v1 = x1.^2;
4  v2 = x2.*sin(v1);
5  f = v1 + v2;
6  % intialize bar variables
7  f_bar = 1.0; v2_bar = 0.0; v1_bar = 0.0;
8  x2_bar = 0.0; x1_bar = 0.0;
9  v2_bar = v2_bar + f_bar; % line 5
10 v1_bar = v1_bar + f_bar; % line 5
11 v1_bar = v1_bar + x2.*cos(v1).*v2_bar; % line 4
12 x2_bar = x2_bar + sin(v1).*v2_bar; % line 4
13 x1_bar = x1_bar + 2.0*x1*v1_bar; % line 3
14 dfdx1 = x1_bar;
15 dfdx2 = x2_bar;
16 end
```

Pros & Cons of Reverse-mode AD

- ✓ no truncation error and no h to worry about!
- ✓ the cost of evaluating the differentiated function is only a small factor more than the original code (approximately twice the cost)
- ✓ produces the entire gradient at once! (cost is virtually independent of the number of design variables)
- ✗ requires access to the source code

Note: you can only get the gradient of one output at a time, therefore...

When Should You Use Forward Mode? Reverse Mode?

Suppose you have a function that takes in n inputs and produces m outputs, and you need the gradient, with respect to all inputs, of all outputs. Generally speaking,

- if $n > m$, use the reverse mode
- if $m > n$, use the forward mode (or complex step)