Automatic Differentiation: CES example

Symbolic derivatives are more accurate than finite differences, but they can be expensive to evaluate without a lot of work. We shall illustrate the points in a simple example.

We write an expression corresponding to four inputs in a CES utility function

In[77]:=
```
vars = {x, y, z, w};
```

In[79]:=
```
utility = Sum[a_i vars[[i]]^σ, {i, 1, 4}]^ϕ
```

Out[79]= $(x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^\phi$

This takes five powers, three additions, and four mutiplications: 12 flops

Compute the gradient

In[81]:=
```
gradutility = Table[D[utility, vars[[i]]], {i, 1, 4}];
gradutility // TableForm
```

Out[82]//TableForm=

$$x^{-1+\sigma}\,\sigma\,\phi\,a_1\,(x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-1+\phi}$$
$$y^{-1+\sigma}\,\sigma\,\phi\,a_2\,(x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-1+\phi}$$
$$z^{-1+\sigma}\,\sigma\,\phi\,a_3\,(x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-1+\phi}$$
$$w^{-1+\sigma}\,\sigma\,\phi\,a_4\,(x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-1+\phi}$$

Each gradient uses six powers, five additions, and eight multiplications: 19x4 = 95 flops
So, computing the four terms in the gradient vector costs EIGHT time the cost of one function evaluation!

However, note that the function and its gradient contains many common elements. Let's use this to reduce the computational burden.

The following term appears in utility and its appearance in each gradient can be avoided by the following substitution

In[83]:= **sub1 = (x^σ a_1 + y^σ a_2 + z^σ a_3 + w^σ a_4) → v1**

Out[83]= $x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4 \to \text{v1}$

In[84]:= **grad = gradf /. sub1**

Out[84]= $\{\text{v1}^{-1+\phi}\, x^{-1+\sigma}\, \sigma\, \phi\, a_1,\ \text{v1}^{-1+\phi}\, y^{-1+\sigma}\, \sigma\, \phi\, a_2,\ \text{v1}^{-1+\phi}\, z^{-1+\sigma}\, \sigma\, \phi\, a_3,\ \text{v1}^{-1+\phi}\, w^{-1+\sigma}\, \sigma\, \phi\, a_4\}$

We now compute the -1+σ, -1+ϕ and $\sigma\,\phi$ terms: 3 new flops
but then substitute them out in grad

In[87]:=
```
sub2 = {-1 + ϕ → ϕm1, -1 + σ → σm1, σ ϕ → σϕ}
```

Out[87]= $\{-1+\phi\to \phi\text{m1},\ -1+\sigma\to \sigma\text{m1},\ \sigma\,\phi\to \sigma\phi\}$

In[88]:=
```
grad = grad //. sub2
```

Out[88]= $\left\{\text{v1}^{\phi\text{m1}}\, x^{\sigma\text{m1}}\, \sigma\phi\, a_1,\ \text{v1}^{\phi\text{m1}}\, y^{\sigma\text{m1}}\, \sigma\phi\, a_2,\ \text{v1}^{\phi\text{m1}}\, z^{\sigma\text{m1}}\, \sigma\phi\, a_3,\ \text{v1}^{\phi\text{m1}}\, w^{\sigma\text{m1}}\, \sigma\phi\, a_4\right\}$

The next substitution requires one power and one multiplication: 2 flops

In[89]:=
```
sub4 = σϕ v1^ϕm1 → v2
```

Out[89]= $\text{v1}^{\phi\text{m1}}\, \sigma\phi \to \text{v2}$

In[90]:=
```
grad = grad /. sub4
```

Out[90]= $\left\{\text{v2}\, x^{\sigma\text{m1}}\, a_1,\ \text{v2}\, y^{\sigma\text{m1}}\, a_2,\ \text{v2}\, z^{\sigma\text{m1}}\, a_3,\ \text{v2}\, w^{\sigma\text{m1}}\, a_4\right\}$

This expression now needs two multiplications and one power per gradient. So the marginal cost of the gradient is 5+4x3=17 flops -- BASICALLY THE SAME COST AS ONE EVALUATION OF THE FUNCTION!

Let's now consider the Hessian. We first define it and then eliminate repetitions due to symmetry

In[91]:=
```
hessf = Table[D[gradf, vars[[i]]], {i, 1, 4}];
hess = hessf // Flatten // Union
```

Out[92]=
$$\{x^{-1+\sigma}\, y^{-1+\sigma}\, \sigma^2\, (-1+\phi)\, \phi\, a_1\, a_2\, (x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-2+\phi},\ x^{-1+\sigma}\, z^{-1+\sigma}\, \sigma^2\, (-1+\phi)\, \phi\, a_1\, a_3\, (x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-2+\phi},$$
$$y^{-1+\sigma}\, z^{-1+\sigma}\, \sigma^2\, (-1+\phi)\, \phi\, a_2\, a_3\, (x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-2+\phi},\ w^{-1+\sigma}\, x^{-1+\sigma}\, \sigma^2\, (-1+\phi)\, \phi\, a_1\, a_4\, (x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-2+\phi},$$
$$w^{-1+\sigma}\, y^{-1+\sigma}\, \sigma^2\, (-1+\phi)\, \phi\, a_2\, a_4\, (x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-2+\phi},\ w^{-1+\sigma}\, z^{-1+\sigma}\, \sigma^2\, (-1+\phi)\, \phi\, a_3\, a_4\, (x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-2+\phi},$$
$$x^{-2+2\sigma}\, \sigma^2\, (-1+\phi)\, \phi\, a_1^2\, (x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-2+\phi} + x^{-2+\sigma}\, (-1+\sigma)\, \sigma\, \phi\, a_1\, (x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-1+\phi},$$
$$y^{-2+2\sigma}\, \sigma^2\, (-1+\phi)\, \phi\, a_2^2\, (x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-2+\phi} + y^{-2+\sigma}\, (-1+\sigma)\, \sigma\, \phi\, a_2\, (x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-1+\phi},$$
$$z^{-2+2\sigma}\, \sigma^2\, (-1+\phi)\, \phi\, a_3^2\, (x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-2+\phi} + z^{-2+\sigma}\, (-1+\sigma)\, \sigma\, \phi\, a_3\, (x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-1+\phi},$$
$$w^{-2+2\sigma}\, \sigma^2\, (-1+\phi)\, \phi\, a_4^2\, (x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-2+\phi} + w^{-2+\sigma}\, (-1+\sigma)\, \sigma\, \phi\, a_4\, (x^\sigma a_1 + y^\sigma a_2 + z^\sigma a_3 + w^\sigma a_4)^{-1+\phi}\}$$

First we use the substitutions we used for the gradient; these require no new flops.

In[99]:=
```
hess = hess /. sub1 /. sub2 /. sub3;
% // TableForm
```

Out[100]//TableForm=

$$\begin{array}{l}
\text{v1}^{-2+\phi}\,x^{\sigma\text{m1}}\,y^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_1\,a_2\\
\text{v1}^{-2+\phi}\,x^{\sigma\text{m1}}\,z^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_1\,a_3\\
\text{v1}^{-2+\phi}\,y^{\sigma\text{m1}}\,z^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_2\,a_3\\
\text{v1}^{-2+\phi}\,w^{\sigma\text{m1}}\,x^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_1\,a_4\\
\text{v1}^{-2+\phi}\,w^{\sigma\text{m1}}\,y^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_2\,a_4\\
\text{v1}^{-2+\phi}\,w^{\sigma\text{m1}}\,z^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_3\,a_4\\
\text{v1}^{\phi\text{m1}}\,x^{-2+\sigma}\,\sigma\text{m1}\,\sigma\phi\,a_1+\text{v1}^{-2+\phi}\,x^{-2+2\,\sigma}\,\sigma^2\,\phi\,\phi\text{m1}\,a_1^2\\
\text{v1}^{\phi\text{m1}}\,y^{-2+\sigma}\,\sigma\text{m1}\,\sigma\phi\,a_2+\text{v1}^{-2+\phi}\,y^{-2+2\,\sigma}\,\sigma^2\,\phi\,\phi\text{m1}\,a_2^2\\
\text{v1}^{\phi\text{m1}}\,z^{-2+\sigma}\,\sigma\text{m1}\,\sigma\phi\,a_3+\text{v1}^{-2+\phi}\,z^{-2+2\,\sigma}\,\sigma^2\,\phi\,\phi\text{m1}\,a_3^2\\
\text{v1}^{\phi\text{m1}}\,w^{-2+\sigma}\,\sigma\text{m1}\,\sigma\phi\,a_4+\text{v1}^{-2+\phi}\,w^{-2+2\,\sigma}\,\sigma^2\,\phi\,\phi\text{m1}\,a_4^2
\end{array}$$

Let's now eliminate repeated terms. The next substitution requires one power and one addition (2 extra flops)

In[101]:= **hess /. v1$^{-2+\sigma 1}$ → v3**

Out[101]=

$$\begin{array}{l}
\{\text{v1}^{-2+\phi}\,x^{\sigma\text{m1}}\,y^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_1\,a_2,\ \text{v1}^{-2+\phi}\,x^{\sigma\text{m1}}\,z^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_1\,a_3,\ \text{v1}^{-2+\phi}\,y^{\sigma\text{m1}}\,z^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_2\,a_3,\\
\ \text{v1}^{-2+\phi}\,w^{\sigma\text{m1}}\,x^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_1\,a_4,\ \text{v1}^{-2+\phi}\,w^{\sigma\text{m1}}\,y^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_2\,a_4,\ \text{v1}^{-2+\phi}\,w^{\sigma\text{m1}}\,z^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_3\,a_4,\\
\ \text{v1}^{\phi\text{m1}}\,x^{-2+\sigma}\,\sigma\text{m1}\,\sigma\phi\,a_1+\text{v1}^{-2+\phi}\,x^{-2+2\,\sigma}\,\sigma^2\,\phi\,\phi\text{m1}\,a_1^2,\ \text{v1}^{\phi\text{m1}}\,y^{-2+\sigma}\,\sigma\text{m1}\,\sigma\phi\,a_2+\text{v1}^{-2+\phi}\,y^{-2+2\,\sigma}\,\sigma^2\,\phi\,\phi\text{m1}\,a_2^2,\\
\ \text{v1}^{\phi\text{m1}}\,z^{-2+\sigma}\,\sigma\text{m1}\,\sigma\phi\,a_3+\text{v1}^{-2+\phi}\,z^{-2+2\,\sigma}\,\sigma^2\,\phi\,\phi\text{m1}\,a_3^2,\ \text{v1}^{\phi\text{m1}}\,w^{-2+\sigma}\,\sigma\text{m1}\,\sigma\phi\,a_4+\text{v1}^{-2+\phi}\,w^{-2+2\,\sigma}\,\sigma^2\,\phi\,\phi\text{m1}\,a_4^2\}
\end{array}$$

one extra addition (cumulative total of 3 flops)

In[102]:=
```
% /. -2 + σ → σm2
```

Out[102]=

$$\begin{array}{l}
\{\text{v1}^{-2+\phi}\,x^{\sigma\text{m1}}\,y^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_1\,a_2,\ \text{v1}^{-2+\phi}\,x^{\sigma\text{m1}}\,z^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_1\,a_3,\ \text{v1}^{-2+\phi}\,y^{\sigma\text{m1}}\,z^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_2\,a_3,\\
\ \text{v1}^{-2+\phi}\,w^{\sigma\text{m1}}\,x^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_1\,a_4,\ \text{v1}^{-2+\phi}\,w^{\sigma\text{m1}}\,y^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_2\,a_4,\ \text{v1}^{-2+\phi}\,w^{\sigma\text{m1}}\,z^{\sigma\text{m1}}\,\sigma^2\,\phi\,\phi\text{m1}\,a_3\,a_4,\\
\ \text{v1}^{\phi\text{m1}}\,x^{\sigma\text{m2}}\,\sigma\text{m1}\,\sigma\phi\,a_1+\text{v1}^{-2+\phi}\,x^{-2+2\,\sigma}\,\sigma^2\,\phi\,\phi\text{m1}\,a_1^2,\ \text{v1}^{\phi\text{m1}}\,y^{\sigma\text{m2}}\,\sigma\text{m1}\,\sigma\phi\,a_2+\text{v1}^{-2+\phi}\,y^{-2+2\,\sigma}\,\sigma^2\,\phi\,\phi\text{m1}\,a_2^2,\\
\ \text{v1}^{\phi\text{m1}}\,z^{\sigma\text{m2}}\,\sigma\text{m1}\,\sigma\phi\,a_3+\text{v1}^{-2+\phi}\,z^{-2+2\,\sigma}\,\sigma^2\,\phi\,\phi\text{m1}\,a_3^2,\ \text{v1}^{\phi\text{m1}}\,w^{\sigma\text{m2}}\,\sigma\text{m1}\,\sigma\phi\,a_4+\text{v1}^{-2+\phi}\,w^{-2+2\,\sigma}\,\sigma^2\,\phi\,\phi\text{m1}\,a_4^2\}
\end{array}$$

one extra addition and one multiplication (total 5 flops)

In[103]:= `% /. -2 + ϕ →  ϕm2`

Out[103]= $\{v1^{\phi m2}\, x^{\sigma m1}\, y^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_2,\ v1^{\phi m2}\, x^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_3,\ v1^{\phi m2}\, y^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_2\, a_3,$
$v1^{\phi m2}\, w^{\sigma m1}\, x^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_4,\ v1^{\phi m2}\, w^{\sigma m1}\, y^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_2\, a_4,\ v1^{\phi m2}\, w^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_3\, a_4,$
$v1^{\phi m1}\, x^{\sigma m2}\, \sigma m1\, \sigma\phi\, a_1 + v1^{\phi m2}\, x^{-2+2\,\sigma}\, \sigma^2\, \phi\, \phi m1\, a_1^2,\ v1^{\phi m1}\, y^{\sigma m2}\, \sigma m1\, \sigma\phi\, a_2 + v1^{\phi m2}\, y^{-2+2\,\sigma}\, \sigma^2\, \phi\, \phi m1\, a_2^2,$
$v1^{\phi m1}\, z^{\sigma m2}\, \sigma m1\, \sigma\phi\, a_3 + v1^{\phi m2}\, z^{-2+2\,\sigma}\, \sigma^2\, \phi\, \phi m1\, a_3^2,\ v1^{\phi m1}\, w^{\sigma m2}\, \sigma m1\, \sigma\phi\, a_4 + v1^{\phi m2}\, w^{-2+2\,\sigma}\, \sigma^2\, \phi\, \phi m1\, a_4^2\}$

In[105]:= `% /. -2 + 2 σ →  σ2m2`

Out[105]= $\{v1^{\phi m2}\, x^{\sigma m1}\, y^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_2,\ v1^{\phi m2}\, x^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_3,\ v1^{\phi m2}\, y^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_2\, a_3,\ v1^{\phi m2}\, w^{\sigma m1}\, x^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_4,$
$v1^{\phi m2}\, w^{\sigma m1}\, y^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_2\, a_4,\ v1^{\phi m2}\, w^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_3\, a_4,\ v1^{\phi m1}\, x^{\sigma m2}\, \sigma m1\, \sigma\phi\, a_1 + v1^{\phi m2}\, x^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_1^2,$
$v1^{\phi m1}\, y^{\sigma m2}\, \sigma m1\, \sigma\phi\, a_2 + v1^{\phi m2}\, y^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_2^2,\ v1^{\phi m1}\, z^{\sigma m2}\, \sigma m1\, \sigma\phi\, a_3 + v1^{\phi m2}\, z^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_3^2,\ v1^{\phi m1}\, w^{\sigma m2}\, \sigma m1\, \sigma\phi\, a_4 + v1^{\phi m2}\, w^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_4^2\}$

one multiplication (total 6 flops)

In[106]:= `% /. σm1 σϕ →  σm1σϕ`

Out[106]= $\{v1^{\phi m2}\, x^{\sigma m1}\, y^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_2,\ v1^{\phi m2}\, x^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_3,\ v1^{\phi m2}\, y^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_2\, a_3,\ v1^{\phi m2}\, w^{\sigma m1}\, x^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_4,$
$v1^{\phi m2}\, w^{\sigma m1}\, y^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_2\, a_4,\ v1^{\phi m2}\, w^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_3\, a_4,\ v1^{\phi m1}\, x^{\sigma m2}\, \sigma m1\sigma\phi\, a_1 + v1^{\phi m2}\, x^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_1^2,$
$v1^{\phi m1}\, y^{\sigma m2}\, \sigma m1\sigma\phi\, a_2 + v1^{\phi m2}\, y^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_2^2,\ v1^{\phi m1}\, z^{\sigma m2}\, \sigma m1\sigma\phi\, a_3 + v1^{\phi m2}\, z^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_3^2,\ v1^{\phi m1}\, w^{\sigma m2}\, \sigma m1\sigma\phi\, a_4 + v1^{\phi m2}\, w^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_4^2\}$

In[107]:= `% /. v1^ϕm2 → v2`

Out[107]= $\{v2\, x^{\sigma m1}\, y^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_2,\ v2\, x^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_3,\ v2\, y^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_2\, a_3,\ v2\, w^{\sigma m1}\, x^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_4,$
$v2\, w^{\sigma m1}\, y^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_2\, a_4,\ v2\, w^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_3\, a_4,\ v1^{\phi m1}\, x^{\sigma m2}\, \sigma m1\sigma\phi\, a_1 + v2\, x^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_1^2,$
$v1^{\phi m1}\, y^{\sigma m2}\, \sigma m1\sigma\phi\, a_2 + v2\, y^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_2^2,\ v1^{\phi m1}\, z^{\sigma m2}\, \sigma m1\sigma\phi\, a_3 + v2\, z^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_3^2,\ v1^{\phi m1}\, w^{\sigma m2}\, \sigma m1\sigma\phi\, a_4 + v2\, w^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_4^2\}$

In[108]:= `% /. v1^ϕm1 → v3`

Out[108]= $\{v2\, x^{\sigma m1}\, y^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_2,\ v2\, x^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_3,\ v2\, y^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_2\, a_3,\ v2\, w^{\sigma m1}\, x^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_1\, a_4,$
$v2\, w^{\sigma m1}\, y^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_2\, a_4,\ v2\, w^{\sigma m1}\, z^{\sigma m1}\, \sigma^2\, \phi\, \phi m1\, a_3\, a_4,\ v3\, x^{\sigma m2}\, \sigma m1\sigma\phi\, a_1 + v2\, x^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_1^2,$
$v3\, y^{\sigma m2}\, \sigma m1\sigma\phi\, a_2 + v2\, y^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_2^2,\ v3\, z^{\sigma m2}\, \sigma m1\sigma\phi\, a_3 + v2\, z^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_3^2,\ v3\, w^{\sigma m2}\, \sigma m1\sigma\phi\, a_4 + v2\, w^{\sigma 2m2}\, \sigma^2\, \phi\, \phi m1\, a_4^2\}$

one multiplication (7 flops)

In[109]:= **% /. $\sigma^2\,\phi\,\phi m1 \to$ SFF // Expand**

Out[109]= $\{SFF\ v2\ x^{\sigma m1}\ y^{\sigma m1}\ a_1\ a_2,\ SFF\ v2\ x^{\sigma m1}\ z^{\sigma m1}\ a_1\ a_3,\ SFF\ v2\ y^{\sigma m1}\ z^{\sigma m1}\ a_2\ a_3,$
$SFF\ v2\ w^{\sigma m1}\ x^{\sigma m1}\ a_1\ a_4,\ SFF\ v2\ w^{\sigma m1}\ y^{\sigma m1}\ a_2\ a_4,\ SFF\ v2\ w^{\sigma m1}\ z^{\sigma m1}\ a_3\ a_4,\ v3\ x^{\sigma m2}\ \sigma m1\sigma\phi\ a_1 + SFF\ v2\ x^{\sigma 2m2}\ a_1^2,$
$v3\ y^{\sigma m2}\ \sigma m1\sigma\phi\ a_2 + SFF\ v2\ y^{\sigma 2m2}\ a_2^2,\ v3\ z^{\sigma m2}\ \sigma m1\sigma\phi\ a_3 + SFF\ v2\ z^{\sigma 2m2}\ a_3^2,\ v3\ w^{\sigma m2}\ \sigma m1\sigma\phi\ a_4 + SFF\ v2\ w^{\sigma 2m2}\ a_4^2\}$

In[112]:= **% /. v2 SFF → v2SFF // Expand**

Out[112]= $\{v2SFF\ x^{\sigma m1}\ y^{\sigma m1}\ a_1\ a_2,\ v2SFF\ x^{\sigma m1}\ z^{\sigma m1}\ a_1\ a_3,\ v2SFF\ y^{\sigma m1}\ z^{\sigma m1}\ a_2\ a_3,\ v2SFF\ w^{\sigma m1}\ x^{\sigma m1}\ a_1\ a_4,\ v2SFF\ w^{\sigma m1}\ y^{\sigma m1}\ a_2\ a_4,\ v2SFF\ w^{\sigma m1}\ z^{\sigma m1}\ a_3\ a_4,$
$v3\ x^{\sigma m2}\ \sigma m1\sigma\phi\ a_1 + v2SFF\ x^{\sigma 2m2}\ a_1^2,\ v3\ y^{\sigma m2}\ \sigma m1\sigma\phi\ a_2 + v2SFF\ y^{\sigma 2m2}\ a_2^2,\ v3\ z^{\sigma m2}\ \sigma m1\sigma\phi\ a_3 + v2SFF\ z^{\sigma 2m2}\ a_3^2,\ v3\ w^{\sigma m2}\ \sigma m1\sigma\phi\ a_4 + v2SFF\ w^{\sigma 2m2}\ a_4^2\}$

This last set requires
4 flops in 6 cases = 24
7 flops in 4 cases = 28

So the total is 12+24+28 = 64 flops, WHICH IS LESS THAN FOUR TIMES THE COST OF THE FUNCTION!