Gradient Descenet

Consider the nonlinear error surface $E(u, v) = (ue^v - 2ve^{-u})^2$. We start at the point (u,v) = (1,1) and minimize this error using gradient descent in the uv space. Use $\eta = 0.1$ (learning rate, not step size).

Gradient Descent (err $\leq 10^{-14}$)

Let's see how many runs it takes until error drops below 10^{-14}

```
(* reset globals *)
Clear[Experiment1]
Clear[e1w, e1η, e1i, e1err]
```

```
Experiment1[] :=
 Module[{
   η = 0.1,
   w = {1, 1},
   err = 1,
   i,
   maxIters = 10 000,
   Ee
  },
  Ee[u_, v_] := (u e^v - 2 v e^-u)^2;
  For[i = 0, i < maxIters ∧ err > 10^-14, i++, (
    w = w - η ∇{u,v} Ee[u, v] /. {u → w[[1]], v → w[[2]]};
    err = Ee[w[[1]], w[[2]]];
   )
  ];
  {w, η, i, err}
 ]
```

```
{e1w, e1η, e1i, e1err} = Experiment1[];
StringForm[
 "Gradient descent results (η=``):\nw=``, i=``, err=``", e1η, e1w, e1i, e1err]
```

```
Gradient descent results (η=0.1`):
w={0.0447363, 0.0239587}, i=10, err=1.2086833939395977`*^-15
```

Comparison with "coordinate descent"

Compare the performance of "coordinate descent." In each iteration, we have two steps along the 2 coordinates. Step 1 is to move only along the u coordinate to reduce the error (assume first-order approximation holds like in gradient descent), and step 2 is to reevaluate and move only along the v coordinate to reduce the error (again, assume first-order approximation holds). Use the same learning rate of $\eta = 0.1$ as we did in gradient descent. Let's check the error E(u,v) after 15 full iterations (30 steps).

```
(* reset globals *)
Clear[Experiment2]
Clear[e2w, e2η, e2i, e2err]
```

```
Experiment2[] :=
 Module[{
    η = 0.1,
    w = {1, 1},
    err = 1,
    i,
    maxIters = 15,
    Ee
   },
   Ee[u_, v_] := (u e^v - 2 v e^-u)^2;
   For[i = 0, i < maxIters ∧ err > 10^-14, i++, (
      w = w - η {(∇{u} Ee[u, v] /. {u → w[[1]], v → w[[2]]})[[1]], 0};
      w = w - η {0, (∇{v} Ee[u, v] /. {u → w[[1]], v → w[[2]]})[[1]]};
      err = Ee[w[[1]], w[[2]]];
    )
   ];
   {w, η, i, err}
 ]
```

```
{e2w, e2η, e2i, e2err} = Experiment2[];
StringForm[
 "Coordinate descent results (η=``):\nw=``, i=``, err=``", e2η, e2w, e2i, e2err]
```

```
Coordinate descent results (η=0.1`):
w={6.29708, -2.85231}, i=15, err=0.13981379199615304`
```