Lagrangian optimization example

When we want to find a minimum or maximum of a function $f(\mathbf{w})$ subject to a constraint $g(\mathbf{w}) = 0$, we can use the Lagrangian approach to solving it.

The Lagrangian is the function $\mathcal{L}(\mathbf{w}, \lambda) = f(\mathbf{w}) + \lambda g(\mathbf{w})$.

If we can find $(\mathbf{w}^*, \lambda^*)$, a stationary point of $\mathcal{L}$, then $\mathbf{w}^*$ is a stationary point of f.

In class we had the simple example $f(\mathbf{w}) = \|w\|^2$ and $g(\mathbf{w}) = w_1 + w_2 - 4$.

By inspection, we can easily see that the solution is $w_1 = w_2 = 2$.

For the analysis, we write the Lagrangian $\mathcal{L}(\mathbf{w}, \lambda) = w_1^2 + w_2^2 + \lambda(w_1 + w_2 - 4)$.

Clearly,

$$\frac{\partial \mathcal{L}}{\partial w_1} = 2w_1 + \lambda$$
$$\frac{\partial \mathcal{L}}{\partial w_2} = 2w_2 + \lambda$$
$$\frac{\partial \mathcal{L}}{\partial \lambda} = w_1 + w_2 - 4.$$

To obtain the stationariy point(s) of this system, we set the partials to 0 and solve. From

$$\lambda = -2w_1 = -2w_2,$$

we obtain $w_1 = w_2$ then the solution $w_1^* = 2, w_2^* = 2, \lambda^* = -4$.

We already knew $(w_1, w_2) = (2, 2)$ by inspection, but now we have confirmed it analytically. We also know it's a minimum by inspection, but to confirm this analytically, we need the Hessian

$$\mathtt{H}_{\mathcal{L}} = \begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial^2 \lambda} & \frac{\partial^2 \mathcal{L}}{\partial \lambda \partial w_1} & \frac{\partial^2 \mathcal{L}}{\partial \lambda \partial w_2} \\ \frac{\partial^2 \mathcal{L}}{\partial w_1 \partial \lambda} & \frac{\partial^2 \mathcal{L}}{\partial^2 w_1} & \frac{\partial^2 \mathcal{L}}{\partial w_1 \partial w_2} \\ \frac{\partial^2 \mathcal{L}}{\partial w_2 \partial \lambda} & \frac{\partial^2 \mathcal{L}}{\partial w_2 \partial w_1} & \frac{\partial^2 \mathcal{L}}{\partial^2 w_2} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

Take a look at the eigenvalues of $\mathtt{H}_{\mathcal{L}}$:

In [4]:
```
import numpy as np

Hl = np.array([[0, 1, 1], [1, 2, 0], [1, 0, 2]])
eigvals, eigvecs = np.linalg.eig(Hl)
print("Eigenvalues: %f, %f, %f" % (eigvals[0], eigvals[1], eigvals[2]))
```

```
Eigenvalues: -0.732051, 2.732051, 2.000000
```

Since we have a mixture of positive and negative eigenvalues (-0.732051, 2.732051, and 2.0), the critical point we've found is a *saddle point* of the Lagrangian. It is actually always the case that the solution of the original optimization is a saddle point of the Lagrangian function.

To determine if the point is a minimum, maximum, or saddle point of the original optimization problem, we consider the determinant of various submatrices of $\mathrm{H}_{\mathcal{L}}$.

For m constraints, we consider the $2m+1$-th up to the nth principal minors of $\mathrm{H}_{\mathcal{L}}$. These are the upper left square submatrices of $\mathrm{H}_{\mathcal{L}}$. In our case, $m=1$ and $n=3=2m+1$, so we have only one principal minor to consider, $\mathrm{H}_{\mathcal{L}}$ itself. We get its determinant:

In [6]:
```
print("Determinant of Hl: %f" % np.linalg.det(Hl))
```

```
Determinant of Hl: -4.000000
```

For a minimum, all of the minors' determinants must have a sign of $(-1)^m$ which means negative for $m=1$. Since this is true, we can conclude that the critical point is a minimum of $f(\mathbf{w})$!

In [ ]: