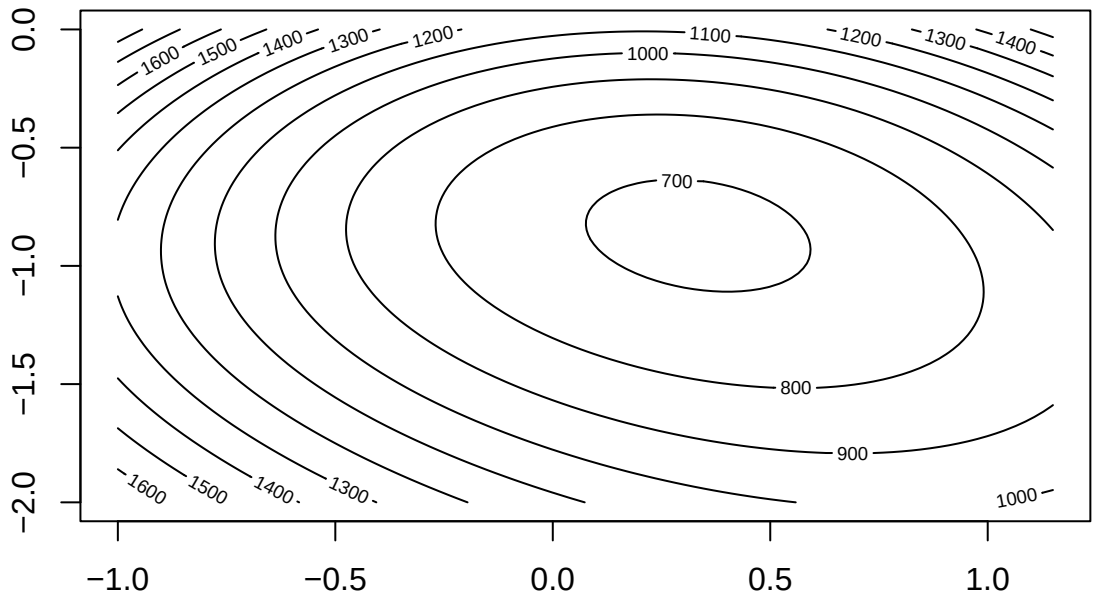
Gradient Competition

Here we will fit the model given by the equation from before via gradient descent but compare the effect of starting values. We generate a contour plot superimposed with the function $-l(\alpha, \beta)$.

```
# ########
# REDACTED
# ########
```

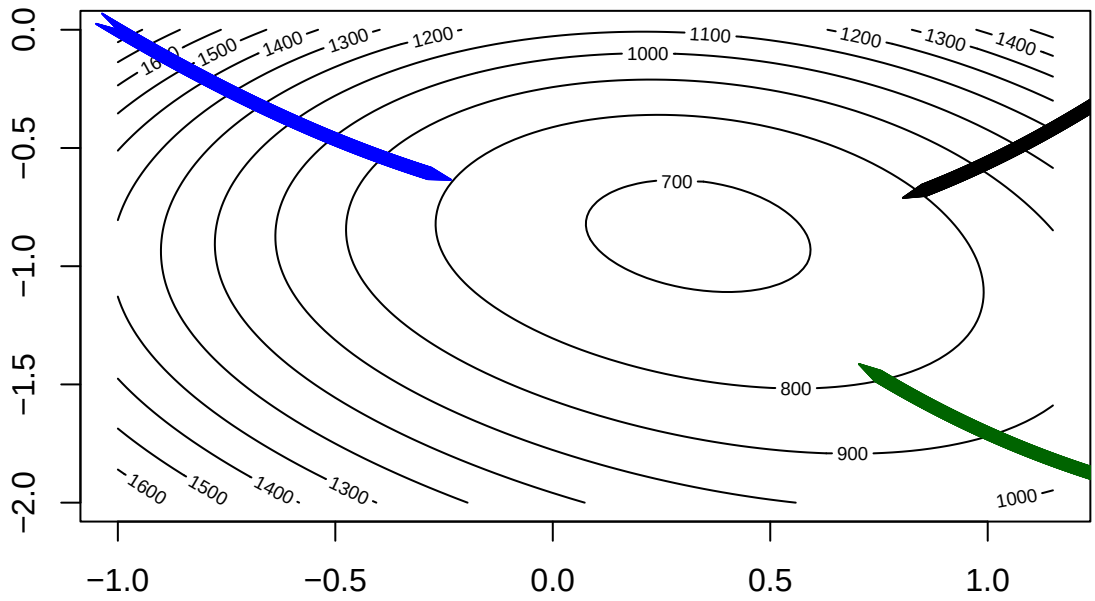


We make a modified gradient descent function to use the unnormalized gradient for the search direction and in addition outputs the sample path.

```
# ########
# REDACTED
# ########
```

We use the initial values $(\alpha^0, \beta^0) = (1.5, 0), (-1, 0), (1.5, -2)$ to perform unnormalized gradient and plot the contour plot and add the three solution paths based on the initial values.

```
# ########
# REDACTED
# ########
```



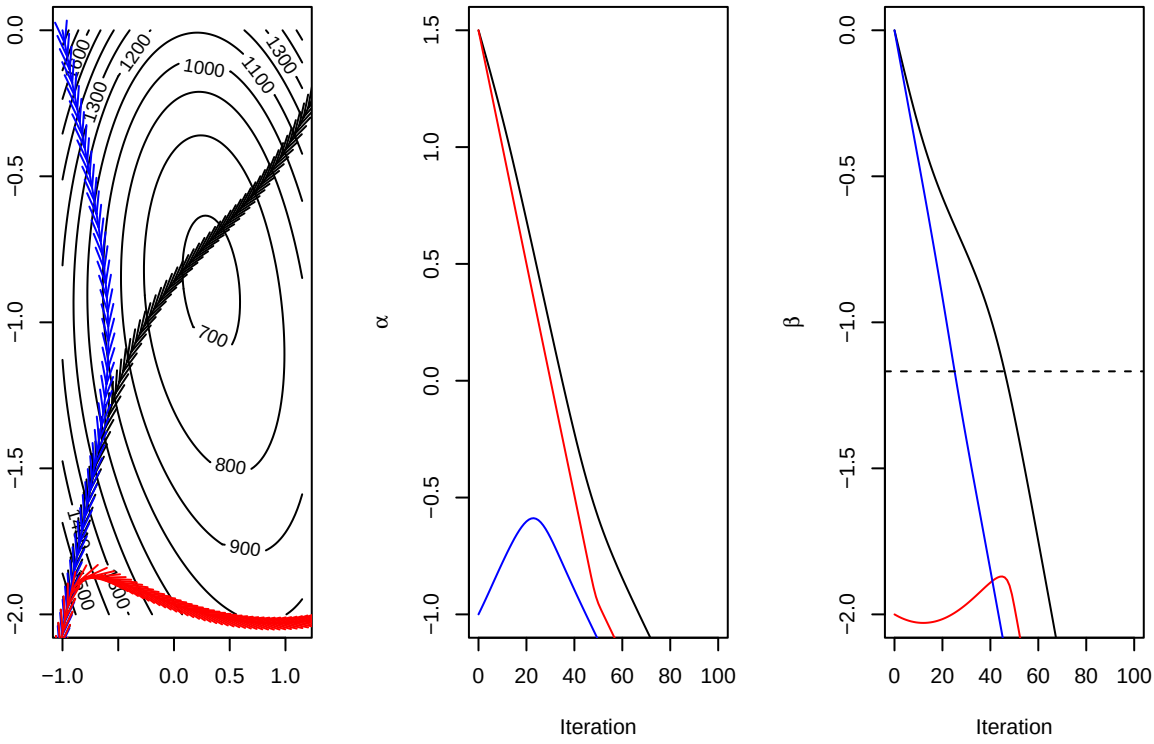

```
##            (1.5, 0)           (-1, 0)            (1.5, -2)
## Theta      "0.805 ,  -0.711"  "-0.234 ,  -0.635" "0.704 ,  -1.414"
## Converged  "FALSE"            "FALSE"            "FALSE"
## Iterations "999"              "999"              "999"
## Fn Value   "779.587521398689" "802.471234378173" "777.586247314438"
```

We create a function called which estimates the gradient of the objective function using a sample.

```
# ########
# REDACTED
# ########
```

We perform random sample stochastic gradient descent using the three initial values $(\alpha^0, \beta^0) = (1.5, 0), (-1, 0), (1.5, -2)$. We construct a figure with a contour plot with the three solution paths based on the initial values, a plot of solution paths for α versus the iteration number and a plot of solution paths for β versus the iteration number.

```
# ########
# REDACTED
# ########
```


0.0
−0.5
−1.0
−1.5
−2.0
−1.0
0.0
0.5
1.0
700
800
900
1000
1100
1200
1300
α
1.5
1.0
0.5
0.0
−0.5
−1.0
0
20
40
60
80
100
Iteration
β
0.0
−0.5
−1.0
−1.5
−2.0
0
20
40
60
80
100
Iteration