

- [1] The smallest loss achieved = x.xxxxxxxxxxxxxx
- [2] The corresponding learning rate = x.xxxxxxxx
- [3] The corresponding $\hat{\beta}$ = [x.xxxxxxx, x.xxxxxxx, x.xxxxxxx, x.xxxxxxx]
Note that the true β = [2.0, 0.5, 3.0, 0.1], and the loss at the true β = 22.627436562328377
- [4] Number of updates used x
- [5] same as the plot in Q
Note that the y-axis tick marks of the plot in Q are scaled and shifted.
A tick mark of 0 represent $22.298851605 = 0 \times (1e-12) + 2.2298851605e+1$
- [6] Same as the plot in Q