

Table 1: Performance Results for LinearRegression\100samplesn5featuresneg0.01Problem

Optimizer	Mean Final Value	Std Dev	Best Value	Worst Value	Mean Func Evals	Success Rate (%)	Mean Time (s)
QQN-Bisection-2	7.15×10^{-2}	1.62×10^{-14}	7.15×10^{-2}	7.15×10^{-2}	103.0	0.0	0.037
QQN-Bisection-1	7.15×10^{-2}	3.15×10^{-14}	7.15×10^{-2}	7.15×10^{-2}	124.5	0.0	0.046
QQN-StrongWolfe	7.15×10^{-2}	6.70×10^{-14}	7.15×10^{-2}	7.15×10^{-2}	110.6	0.0	0.039
QQN-Backtracking	7.15×10^{-2}	3.44×10^{-14}	7.15×10^{-2}	7.15×10^{-2}	83.0	0.0	0.026
QQN-CubicQuadraticInterpolation	7.15×10^{-2}	2.72×10^{-10}	7.15×10^{-2}	7.15×10^{-2}	131.7	0.0	0.043
QQN-GoldenSection	7.15×10^{-2}	7.23×10^{-10}	7.15×10^{-2}	7.15×10^{-2}	326.5	0.0	0.060
QQN-MoreThuente	7.15×10^{-2}	5.59×10^{-10}	7.15×10^{-2}	7.15×10^{-2}	91.2	0.0	0.029
L-BFGS-Conservative	7.15×10^{-2}	1.66×10^{-8}	7.15×10^{-2}	7.15×10^{-2}	553.2	0.0	0.167
GD-WeightDecay	7.15×10^{-2}	1.85×10^{-10}	7.15×10^{-2}	7.15×10^{-2}	165.3	0.0	0.084
GD	7.16×10^{-2}	5.10×10^{-7}	7.15×10^{-2}	7.16×10^{-2}	335.0	0.0	0.170
L-BFGS-Aggressive	7.20×10^{-2}	1.90×10^{-3}	7.15×10^{-2}	8.04×10^{-2}	309.3	0.0	0.066
L-BFGS	8.02×10^{-2}	3.22×10^{-2}	7.15×10^{-2}	2.20×10^{-1}	239.8	0.0	0.061
Adam-Fast	2.10×10^{-1}	5.29×10^{-2}	7.27×10^{-2}	2.53×10^{-1}	88.3	0.0	0.029
GD-Nesterov	9.04×10^{-1}	3.38×10^{-2}	8.18×10^{-1}	9.49×10^{-1}	23.8	0.0	0.012
GD-Momentum	1.38	4.50×10^{-2}	1.31	1.49	24.1	0.0	0.012
Adam-WeightDecay	7.28	5.46×10^{-1}	6.17	7.93	502.0	0.0	0.166
Adam	1.31×10^1	9.42×10^{-1}	1.17×10^1	1.45×10^1	502.0	0.0	0.165
Adam-AMSGrad	1.34×10^1	8.32×10^{-1}	1.17×10^1	1.52×10^1	502.0	0.0	0.166
Trust Region-Conservative	1.57×10^1	1.19	1.32×10^1	1.79×10^1	602.0	0.0	0.155
Trust Region-Standard	1.21×10^2	4.92×10^1	3.26×10^{-1}	1.70×10^2	76.8	0.0	0.020
Trust Region-Adaptive	1.81×10^3	1.39×10^2	1.54×10^3	1.96×10^3	232.8	0.0	0.060