

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-1	7.15×10^{-2}	2.31×10^{-14}	7.15×10^{-2}	7.15×10^{-2}	124.5	0.0	0.049
QQN-Bisection-2	7.15×10^{-2}	2.46×10^{-14}	7.15×10^{-2}	7.15×10^{-2}	103.0	0.0	0.039
QQN-StrongWolfe	7.15×10^{-2}	4.54×10^{-14}	7.15×10^{-2}	7.15×10^{-2}	111.3	0.0	0.042
QQN-Backtracking	7.15×10^{-2}	2.95×10^{-14}	7.15×10^{-2}	7.15×10^{-2}	83.0	0.0	0.027
QQN-CubicQuadraticInterpolation	7.15×10^{-2}	4.21×10^{-10}	7.15×10^{-2}	7.15×10^{-2}	133.6	0.0	0.046
QQN-MoreThunten	7.15×10^{-2}	4.48×10^{-10}	7.15×10^{-2}	7.15×10^{-2}	89.7	0.0	0.031
QQN-GoldenSection	7.15×10^{-2}	1.07×10^{-9}	7.15×10^{-2}	7.15×10^{-2}	328.1	0.0	0.064
L-BFGS-Conservative	7.15×10^{-2}	1.06×10^{-8}	7.15×10^{-2}	7.15×10^{-2}	557.0	0.0	0.172
GD-WeightDecay	7.15×10^{-2}	2.52×10^{-10}	7.15×10^{-2}	7.15×10^{-2}	164.5	0.0	0.085
GD	7.16×10^{-2}	4.73×10^{-7}	7.15×10^{-2}	7.16×10^{-2}	335.0	0.0	0.171
L-BFGS-Aggressive	7.18×10^{-2}	1.51×10^{-3}	7.15×10^{-2}	8.12×10^{-2}	320.3	0.0	0.070
L-BFGS	1.38×10^{-1}	1.61×10^{-1}	7.15×10^{-2}	7.27×10^{-1}	187.8	0.0	0.050
Adam-Fast	2.18×10^{-1}	4.29×10^{-2}	7.26×10^{-2}	2.65×10^{-1}	83.8	0.0	0.028
GD-Nesterov	9.14×10^{-1}	3.26×10^{-2}	8.32×10^{-1}	9.58×10^{-1}	23.9	0.0	0.012
GD-Momentum	1.36	3.40×10^{-2}	1.31	1.48	24.1	0.0	0.012
Adam-WeightDecay	6.82	4.15×10^{-1}	6.16	7.76	502.0	0.0	0.167
Adam-AMSGrad	1.35×10^1	9.74×10^{-1}	1.19×10^1	1.51×10^1	502.0	0.0	0.168
Adam	1.37×10^1	8.20×10^{-1}	1.20×10^1	1.50×10^1	502.0	0.0	0.167
Trust Region-Conservative	1.57×10^1	1.16	1.34×10^1	1.79×10^1	602.0	0.0	0.156
Trust Region-Standard	8.52×10^1	6.20×10^1	2.87×10^{-1}	1.69×10^2	73.0	0.0	0.019
Trust Region-Adaptive	1.39×10^3	8.09×10^2	1.30×10^{-1}	1.96×10^3	235.4	0.0	0.062