

Table 1: Performance Results for LinearRegression\200samplesn10featuresnreg0.01Problem

Optimizer	Mean Final Value	Std Dev	Best Value	Worst Value	Mean Func Evals	Success Rate (%)	Mean Time (s)
QQN-Bisection-1	4.82×10^{-1}	4.61×10^{-14}	4.82×10^{-1}	4.82×10^{-1}	143.8	100.0	0.100
QQN-Bisection-2	4.82×10^{-1}	4.75×10^{-14}	4.82×10^{-1}	4.82×10^{-1}	121.7	100.0	0.080
QQN-MoreThuente	4.82×10^{-1}	5.41×10^{-14}	4.82×10^{-1}	4.82×10^{-1}	138.4	100.0	0.086
QQN-GoldenSection	4.82×10^{-1}	6.59×10^{-14}	4.82×10^{-1}	4.82×10^{-1}	362.1	100.0	0.124
QQN-StrongWolfe	4.82×10^{-1}	8.18×10^{-15}	4.82×10^{-1}	4.82×10^{-1}	126.0	100.0	0.086
QQN-Backtracking	4.82×10^{-1}	9.25×10^{-15}	4.82×10^{-1}	4.82×10^{-1}	104.0	100.0	0.062
QQN-CubicQuadraticInterpolation	4.82×10^{-1}	1.40×10^{-9}	4.82×10^{-1}	4.82×10^{-1}	161.2	100.0	0.108
GD-WeightDecay	4.82×10^{-1}	2.82×10^{-10}	4.82×10^{-1}	4.82×10^{-1}	226.4	100.0	0.213
L-BFGS	4.89×10^{-1}	3.05×10^{-2}	4.82×10^{-1}	6.22×10^{-1}	181.3	100.0	0.090
L-BFGS-Aggressive	5.15×10^{-1}	6.00×10^{-2}	4.82×10^{-1}	7.19×10^{-1}	486.6	67.5	0.192
Adam-Fast	5.05×10^{-1}	3.82×10^{-2}	4.82×10^{-1}	5.80×10^{-1}	217.3	20.0	0.132
L-BFGS-Conservative	4.82×10^{-1}	8.53×10^{-7}	4.82×10^{-1}	4.82×10^{-1}	557.0	0.0	0.311
GD	4.82×10^{-1}	1.84×10^{-5}	4.82×10^{-1}	4.82×10^{-1}	335.0	0.0	0.313
GD-Nesterov	1.80	1.31×10^{-2}	1.78	1.82	38.0	0.0	0.035
GD-Momentum	2.46	2.78×10^{-2}	2.40	2.51	38.0	0.0	0.035
Trust Region-Adaptive	6.63×10^1	3.07	6.02×10^1	7.04×10^1	602.0	0.0	0.288
Adam-WeightDecay	7.49×10^1	2.00	7.12×10^1	7.89×10^1	502.0	0.0	0.304
Adam	1.00×10^2	2.84	9.57×10^1	1.06×10^2	502.0	0.0	0.304
Adam-AMSGrad	1.01×10^2	2.42	9.64×10^1	1.06×10^2	502.0	0.0	0.305
Trust Region-Conservative	1.15×10^2	3.08	1.09×10^2	1.20×10^2	602.0	0.0	0.289
Trust Region-Standard	2.77×10^3	4.39×10^3	7.10×10^{-1}	1.27×10^4	368.6	0.0	0.177