

Table 1: Performance Results for NoisySphere\10Dnigma0.01Problem

Optimizer	Mean Final Value	Std Dev	Best Value	Worst Value	Mean Func Evals	Success Rate (%)	Mean Time (s)
GD-WeightDecay	$1.01 \times 10^1$	2.18	2.03	$1.27 \times 10^1$	15.1	5.0	0.005
L-BFGS	$3.12 \times 10^1$	$1.27 \times 10^1$	$1.37 \times 10^1$	$5.43 \times 10^1$	66.5	5.0	0.005
L-BFGS-Conservative	5.56	1.17	3.61	8.49	829.6	0.0	0.041
QQN-StrongWolfe	7.97	2.20	3.70	$1.08 \times 10^1$	41.5	0.0	0.004
QQN-Bisection-1	8.40	1.84	3.79	$1.23 \times 10^1$	139.5	0.0	0.038
QQN-Backtracking	8.96	1.76	5.06	$1.11 \times 10^1$	213.6	0.0	0.012
QQN-GoldenSection	9.49	1.68	2.52	$1.11 \times 10^1$	29.5	0.0	0.002
QQN-MoreThuente	9.65	$8.12 \times 10^{-1}$	7.67	$1.10 \times 10^1$	24.4	0.0	0.005
Trust Region-Adaptive	9.81	$7.42 \times 10^{-1}$	8.67	$1.11 \times 10^1$	5.0	0.0	0.001
Adam	9.91	$7.60 \times 10^{-1}$	8.61	$1.13 \times 10^1$	29.2	0.0	0.005
Adam-WeightDecay	9.94	$6.33 \times 10^{-1}$	8.59	$1.11 \times 10^1$	31.5	0.0	0.006
Adam-AMSGrad	$1.00 \times 10^1$	$6.98 \times 10^{-1}$	9.06	$1.16 \times 10^1$	29.3	0.0	0.005
L-BFGS-Aggressive	$1.00 \times 10^1$	$6.29 \times 10^{-1}$	8.87	$1.10 \times 10^1$	9.6	0.0	0.001
Trust Region-Conservative	$1.00 \times 10^1$	$6.46 \times 10^{-1}$	8.49	$1.10 \times 10^1$	6.0	0.0	0.001
Trust Region-Standard	$1.01 \times 10^1$	$5.44 \times 10^{-1}$	9.00	$1.13 \times 10^1$	5.0	0.0	0.001
QQN-CubicQuadraticInterpolation	$1.01 \times 10^1$	$8.74 \times 10^{-1}$	7.66	$1.14 \times 10^1$	25.2	0.0	0.005
GD	$1.02 \times 10^1$	$9.45 \times 10^{-1}$	7.75	$1.18 \times 10^1$	13.9	0.0	0.004
GD-Momentum	$1.03 \times 10^1$	1.30	7.58	$1.32 \times 10^1$	13.5	0.0	0.004
GD-Nesterov	$1.11 \times 10^1$	$8.17 \times 10^{-1}$	$1.01 \times 10^1$	$1.29 \times 10^1$	11.7	0.0	0.003
Adam-Fast	$1.14 \times 10^1$	1.56	6.08	$1.48 \times 10^1$	18.9	0.0	0.003