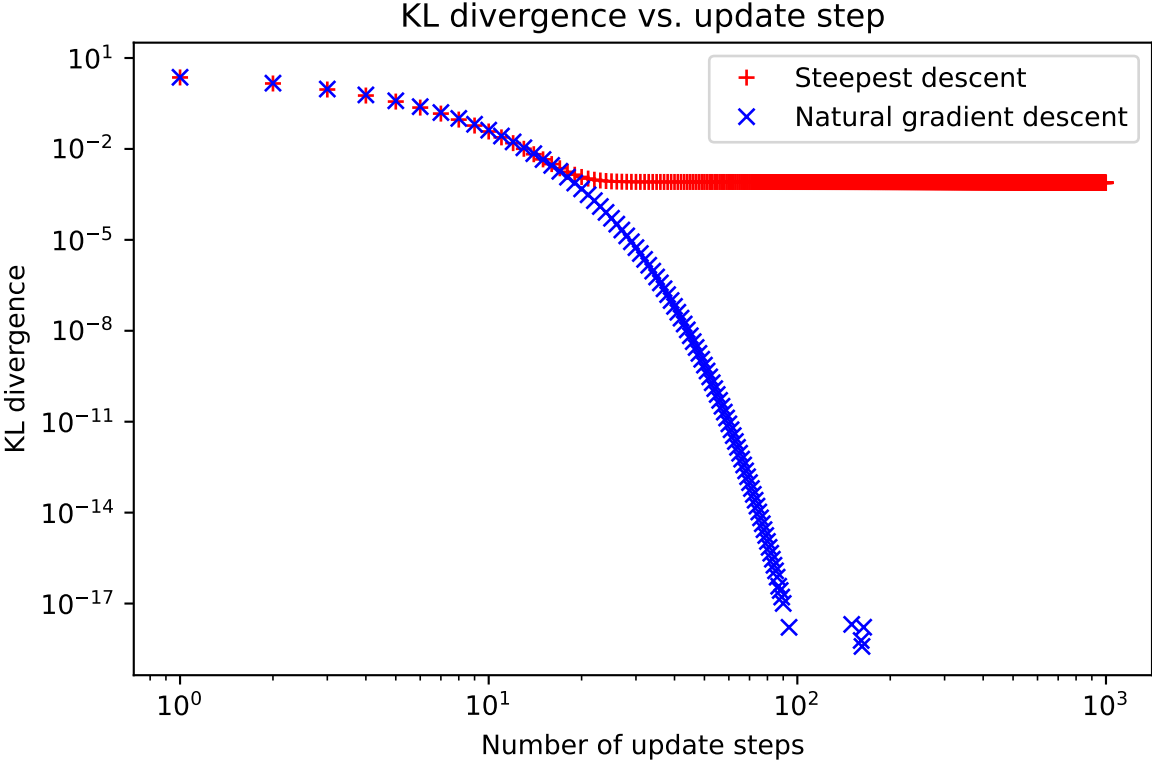
KL divergence vs. update step
Steepest descent
Natural gradient descent
KL divergence
Number of update steps
10^{1}
10^{-2}
10^{-5}
10^{-8}
10^{-11}
10^{-14}
10^{-17}
10^{0}
10^{1}
10^{2}
10^{3}