

The equation I chose to use was the distance formula/pythagorean, but with a matrix in it instead of just normal with two sides of a triangle. I'm not entirely sure what this would end up representing but it was interesting and worked out for the optimization problem. With the learning rates, I found that anything under 0.1 basically modulates so slowly that it just ends up minimizing before it can ever form that sine wave shape. That being said they end up taking forever to learn and didn't minimize until about like 5 times as long as 0.1 for 0.01. For me and my equation 0.25 and 0.5 seemed to be the most stable while 0.75 modulated a bit too quickly and ended up taking longer to minimize than 0.5.