MMSE Based Channel Estimator

Zhao-Jie, Luo

janny00kevin@gmail.com

Advisor: Professor Carrson C. Fung

National Yang Ming Chiao Tung University

Mar. 14, 2024

Reference

1. Dmytro Ivakhnenkov, and Carrson C. Fung, "Model-Driven Neural Network Based MIMO Channel Estimator"
2. M. Eisen, C. Zhang, L.F.O. Chamon, D.D. Lee and A. Ribeiro, "Learning optimal power allocations in wireless systems," *IEEE Trans. on Signal Processing*, vol. 67(10), pp. 2775-2790, May 2019.
3. OpenAI Spinning Up introduction to RL Part 3: Intro to Policy Optimization

Problem Formulation

Our objective is to minimize the expected mean square error:

$$\min_{\hat{\mathbf{h}}} \mathbb{E}_{\mathbf{y},\mathbf{h}} \left[\left\| \mathbf{h} - \hat{\mathbf{h}} \right\|_2^2 \right]$$

and it can be written in epigraph form as:

$$\begin{aligned} &\min_{t,\mathbf{h}} t \\ &\quad s.t.\ \mathbb{E}_{\mathbf{y},\mathbf{h}} \left[\left\| \mathbf{h} - \hat{\mathbf{h}} \right\|_2^2 \right] \end{aligned}$$

Primal-Dual Optimization Mehthod (1)

We use parameterize channel estimator so that $\hat{\mathbf{h}} = \phi(\mathbf{y}; \boldsymbol{\theta})$, with $\boldsymbol{\theta}$ denoting the parameters of the neural network.

Then the Lagrangian function of (15) can be written as

$$\begin{aligned}\mathcal{L}\left(\hat{\mathbf{h}}, t, \lambda\right) &= t + \lambda\left(\mathbb{E}_{\mathbf{y},\mathbf{h}}\left[\left\|\mathbf{h} - \hat{\mathbf{h}}\right\|_2^2\right] - t\right) \\ &= t + \lambda\left(\mathbb{E}_{\mathbf{y},\mathbf{h}}\left[\|\mathbf{h} - \phi(\mathbf{y}; \boldsymbol{\theta})\|_2^2\right] - t\right)\end{aligned}$$

Primal-Dual Optimization Mehthod (2)

It is uncertain whether or not the duality gap equals zero.
However, the stationary point of $\mathcal{L}\left(\hat{\mathbf{h}}, t, \lambda\right)$ can be found via the KKT conditions by solving for the primal and dual variables alternately using gradient descent and ascent, respectively:

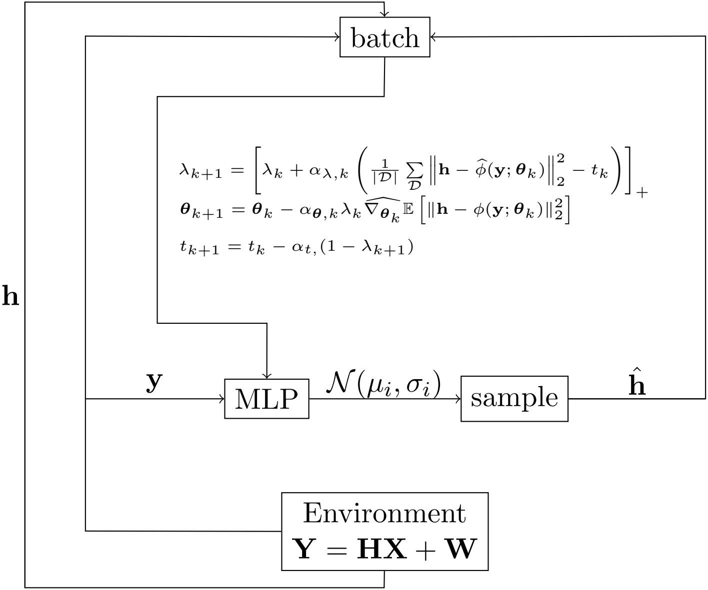
$$
\begin{aligned}
\boldsymbol{\theta}_{k+1} &= \boldsymbol{\theta}_k - \alpha_{\boldsymbol{\theta},k}\lambda_k \nabla_{\boldsymbol{\theta}_k} \mathbb{E}\left[\|\mathbf{h} - \phi(\mathbf{y}; \boldsymbol{\theta}_k)\|_2^2\right] \\
t_{k+1} &= t_k - \alpha_{t,k}(1 - \lambda_k) \\
\lambda_{k+1} &= \left[\lambda_k + \alpha_{\lambda,k}\left(\mathbb{E}\left[\|\mathbf{h} - \phi(\mathbf{y}; \boldsymbol{\theta}_{k+1})\|_2^2\right] - t_{k+1}\right)\right]_+
\end{aligned}
$$

Policy Gradient

Follow the same way as described in Eisen's paper:

$$\widehat{\nabla_{\boldsymbol{\theta}}}\mathbb{E}\left[\|\mathbf{h}-\phi(\mathbf{y};\boldsymbol{\theta})\|_2^2\right]=\left\|\mathbf{h}-\widehat{\phi}(\mathbf{y};\boldsymbol{\theta})\right\|_2^2\nabla_{\boldsymbol{\theta}}\log\pi_{\mathbf{y},\boldsymbol{\theta}}\left(\widehat{\phi}(\mathbf{y};\boldsymbol{\theta})\right)$$

Experiment Diagram



Simulation Result

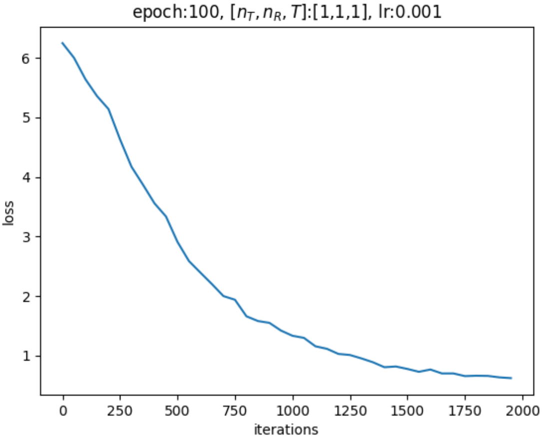


Figure: hidden layer sizes: [64,32], $\mu_H = 5, \sigma_H = 0.2, \mu_W = 0, \sigma_W = 0.1$

Simulation Result

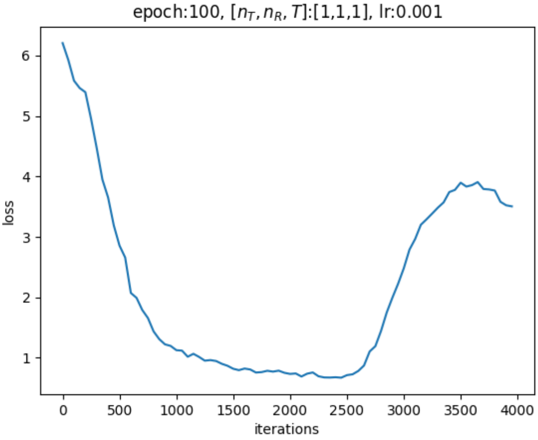


Figure: increase the iterations from 2000 to 4000