Estimation of Channel Distribution Functions using a Neural Network

Peter Hartig

March 13, 2020

Outline

The Channel State

- Observations are made of some channel in a point-to-point communication system.
- For each observation, this channel takes on a state $s[k] \in \mathcal{S}$.
- The true state $s[k]$ is hidden by the addition of noise to an observation $y[k]$.

Sampling Channel State

Over many observations, the sequence $\mathbf{y}$ corresponds to a sequence of channel states $\mathbf{s} \in S^N$

$$\boxed{\text{Transmitter: } \mathbf{x}} \longrightarrow \boxed{\text{Channel: } \mathbf{s}} \longrightarrow \boxed{\text{Receiver: } \mathbf{y}}$$

Sampling Channel State

Over many observations, the sequence $\mathbf{y}$ corresponds to a sequence of channel states $\mathbf{s} \in S^N$

$$\boxed{\text{Transmitter: } \mathbf{x}} \longrightarrow \boxed{\text{Channel: } \mathbf{s}} \longrightarrow \boxed{\text{Receiver: } \mathbf{y}}$$

For a channel represented by an LTI system, the state is determined entirely by the transmitted information $\mathbf{x}$.

Estimating the True Channel State

Goal:

We attempt to estimate the true, hidden, sequence of channel states, $\mathbf{s}$, based the sequence of samples $\mathbf{y}$.

Note

We assume that we known how many states the channel $|S|$

Outline

MAP Sequence Detection

$$\underset{\mathbf{s} \in S^N}{\text{maximize}}\ p(\mathbf{s}|\mathbf{y}).$$

MAP Sequence Detection

$$\underset{\mathbf{s}\in S^N}{\text{maximize}}\ p(\mathbf{s}|\mathbf{y}).$$

Using Bayes' theorem

$$p(\mathbf{s}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{s})p(\mathbf{s})}{p(\mathbf{y})}$$

Noting that $p(\mathbf{y})$ can be ignored

MAP Sequence Detection

$$\underset{\mathbf{s}\in S^N}{\text{maximize}}\; p(\mathbf{s}|\mathbf{y}).$$

Using Bayes' theorem

$$p(\mathbf{s}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{s})p(\mathbf{s})}{p(\mathbf{y})}$$

Noting that $p(\mathbf{y})$ can be ignored

$$\underset{\mathbf{s}\in S^N}{\text{maximize}}\; p(\mathbf{y}|\mathbf{s})p(\mathbf{s}) \tag{1}$$

Example with LTI channel

$$\underset{\mathbf{s} \in S^N}{\text{maximize}}\ p(\mathbf{y}|\mathbf{s})p(\mathbf{s})$$

Example with LTI channel

$$\underset{\mathbf{s}\in S^N}{\text{maximize}}\; p(\mathbf{y}|\mathbf{s})p(\mathbf{s})$$

Assuming

$$p(\mathbf{y}|\mathbf{s}) = \prod_{k=0}^{N-1} p(y[k]|\mathbf{s})$$

Example with LTI channel

$$\underset{\mathbf{s}\in S^N}{\text{maximize}}\; p(\mathbf{y}|\mathbf{s})p(\mathbf{s})$$

Assuming

$$p(\mathbf{y}|\mathbf{s}) = \prod_{k=0}^{N-1} p(y[k]|\mathbf{s})$$

and

$$p(y[k]|\mathbf{s}) = p(y[k]|s[k])$$

Example with LTI channel

$$\underset{\mathbf{s}\in S^N}{\text{maximize}}\ p(\mathbf{y}|\mathbf{s})p(\mathbf{s})$$

Assuming

$$p(\mathbf{y}|\mathbf{s}) = \prod_{k=0}^{N-1} p(y[k]|\mathbf{s})$$

and

$$p(y[k]|\mathbf{s}) = p(y[k]|s[k])$$

this simplifies to

$$\underset{\mathbf{s}\in S^N}{\text{maximize}}\ \prod_{k=0}^{N-1} p(y[k]|s[k])p(\mathbf{s}).$$

Example with LTI channel - Continued

For the LTI channel

$$p(\mathbf{s}) = p(s[N]|s[N-1]\ldots s[0])p(s[N-1]|s[N-2]\ldots s[0])\ldots p(s[1]|s[0])p(s[0])$$

describes the consistency of transmitted symbols implied by the state sequence. The channel states of the LTI channel satisfy the Markov property

$$p(s[N]|s[N-1]\ldots s[0]) = p(s[N]|s[N-1]]).$$

Example with LTI channel - Continued

With these assumptions,

$$\underset{\mathbf{s}\in S^N}{\text{maximize}} \prod_{k=0}^{N-1} p(y[k]|s[k])p(\mathbf{s})$$

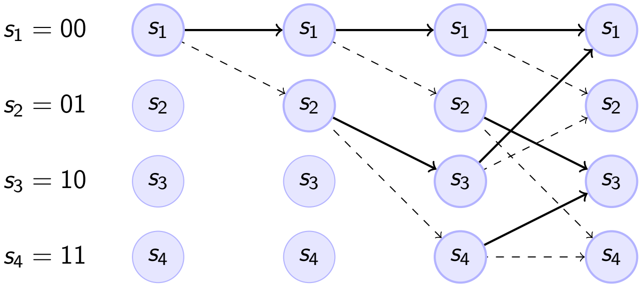
is equivalent to

$$\underset{\mathbf{s}\in S^N}{\text{minimize}} \sum_{k=0}^{N-1} -log(p(y[k]|s[k])p(s[k]|s[k-1])).$$

For the LTI channel $p(s[k]|s[k-1])$ is 0 if states contradict transmission sequence, otherwise this term is constant.

Viterbi Algorithm

$$\underset{\mathbf{s}\in S^N}{\text{minimize}} \sum_{k=0}^{N-1} -log(p(y[k]|s[k])p(s[k]|s[k-1])).$$



Example with channel impulse response length 2 and constellation size 2.

Outline

Decomposing Terms in the Viterbi Algorithm

The individual terms in

$$\underset{\mathbf{s}\in S^N}{\text{minimize}} \sum_{k=0}^{N-1} -log(p(y[k]|s[k])p(s[k]|s[k-1])).$$

can be rewritten

$$p(y[k]|s[k])p(s[k]|s[k-1]) = \frac{p(s[k]|y[k])p(y[k])}{p(s[k])}p(s[k]|s[k-1]).$$

Decomposing Terms in the Viterbi Algorithm

$$\frac{p(s[k]|y[k])p(y[k])}{p(s[k])}$$

Neural Network Component

Mixture Model Component

State Only Components

Outline

State Redundancy

Exploiting State Redundancy

Don't go into details about how this is solved.

Outline

Detection Performance

Without ISI

With ISI

Detection Performance

Reduced Training data (100 vs. 1000 symbols)

Outline

Next Steps

Discuss how this can be applied to other factor graph related algorithms.
Testing on more complicated channels

- Improve decoding performance with neural net.
- Apply to a sampled molecular communications channel.
  - Estimate matched filter
- Generate training data for molecular communications channel and test "transfer learning" to real data.

Thank You.

Questions or Comments?