# Conditional Random Field

## Lateral Weights

**Normal Softmax Function from ANN**

**Regular Classification:**

**V is a matrix connecting yk,yk+1, tells us how likely we see both yk,yk+1 together**

**Joint distribution y given dataset X**

**Sequence classification with linear chain:**

## Context Window

Context window of radius 1 (version 1):

Context window of radius 1 (version 2):

**Different neural networks (-1,0,1) for each set of X values (with neighbours +-1)**

**Combine neighbouring datasets X and feed into the 1 neural network as 1 input**

## Unary and pairwise log-factors

-unary log-factors (depends on 1 label, in this case, that 1 label y=k

or

-pairwise log-factors

Then we have (context window of radius 1 example):

## Inference

alternatively, you can treat this as

based on step 2’s formula.

Treat =1 to obtain the same formula.

**(AKA all possibilities, complexity = O(CK), exponential run time)**

## Belief propagation

this is alpha, the forward pass

### Forward pass

1. Initialize the following:
2. For k=2 toK-1 for all values of :

For last layer,

1. Fill up the α table (time complexity = O(KC2) (for this example, we have 3 classes)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **k=1** | **k=2** | **k=3** | **…** | **k=K-1** | **k=K** |
|  |  |  |  | … |  |  |
|  |  |  |  | … |  |  |
|  |  |  |  | … |  |  |

**Z(X)**

+

+

+

+

### Backward pass

1. Initialize the following:
2. for k=K-1 to 2, for all values of

For last layer:

=

1. Fill up the β table (time complexity = O(KC2) (for this example, we have 3 classes)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **k=1** | **k=2** | **…** | **k=K-2** | **k=K-1** | **k=K** |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  | … |  |  |  |

**Z(X)**

+

+

### Log space implementation

\*for numerical stability

solve log expression as follows:

### Marginals computation

α/β can be used for calculating marginals:

V1: **)**

V2: **)**

Joint marginals:

## Prediction

### Option 1

-At each position ***k***, pick label ***yk*** with the highest marginal probability **p(yk|X)**

(minimising error)

=

=

# End