**1 Auto Encoders**

Using activation function , as thus its derivative is as the following,

and the activation at output layer, so that building a network to minimize the loss

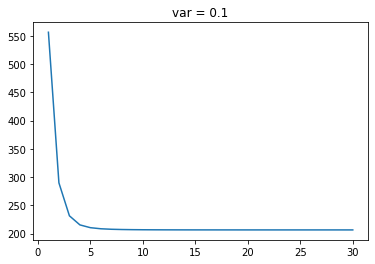
Calculate :

Let be inputs, be hidden layer, be output layer, for .

Add to help calculation. Let be the gradients, where

*For , train an auto encoder on the data, and plot the final loss as a function of*

* *Plot the final loss as a function of for initial*
* *How does this change as changes ?*

Shape, square

Description automatically generatedShape, square

Description automatically generatedShape

Description automatically generatedA picture containing chart

Description automatically generatedChart

Description automatically generated

Around , the variation become small, which implies that the dimension of this set is around .

For variances small, , the losses are decreasing as increases. And the curve of is relatively smooth. While for larger , larger makes the fluctuation larger, so as the value of losses.

The reason might be the following: the data () itself is depends on each other. The smaller variance implies a stronger connection, the data points is closer to each other. Then, using the auto-encoder, it is easier to generate the gradients that closer to “True” results. However, the larger variance makes the connections weaker, which also reduces the dependency in each other, and the dimension become larger. Thus, it is possible that make the nodes more depends on the constant or other noises that should not be use in calculation. Therefore, larger make the auto-encoder less accurate.

**2 PCA**

For the initial , the dimensionality of the data set is . This result is robust as get the same result doing again on a different data set.

A picture containing text, clock

Description automatically generated

For all we choose, there are eigenvalues corresponding to eigenvectors.

Shape, rectangle

Description automatically generatedShape, rectangle

Description automatically generated

For small variance, i.e., , it is reasonable to see the dimensionality as as mainly depends on . While the variance is around , the result is variance between 4 and 5. And for larger variance, the dimensionality is .

**3 Correlation Graphs**

The following is dependency graph, connecting each feature to the other two / three / four features, where the same color represents the “True” dependency structure should be.

Here, using 2 features is robust as getting the same result when doing it again on a different data set.

However, when using larger number of other features, it becomes less robust.

Table

Description automatically generated

It is possible to reconstruct the true dependency graph, but the linear regression should be more accurate. Here is my trying.

A picture containing diagram

Description automatically generated Diagram

Description automatically generated with medium confidence A picture containing text

Description automatically generated

Using 2 features should be enough.

Table

Description automatically generated

For different variance, my trying is the following

start var = 0.01

X 1

-> X 27

-> X 22

-> X 21

-> X 4

-> X 10

-> X 17

-> X 16

-> X 29

-> X 12

-> X 15

X 1

-> X 10

-> X 17

-> X 16

-> X 29

-> X 22

-> X 21

-> X 4

-> X 6

-> X 15

-> X 24

X 1

-> X 17

-> X 16

-> X 29

-> X 22

-> X 21

-> X 4

-> X 10

-> X 13

-> X 9

-> X 19

start var = 0.51

X 1

-> X 16

-> X 13

-> X 21

-> X 26

-> X 15

-> X 12

-> X 18

-> X 9

-> X 6

-> X 7

X 1

-> X 27

-> X 16

-> X 13

-> X 21

-> X 26

-> X 15

-> X 12

-> X 18

-> X 9

-> X 6

X 1

-> X 18

-> X 9

-> X 6

-> X 7

-> X 10

-> X 4

-> X 3

-> X 17

-> X 11

-> X 20

start var = 1.01

X 1

-> X 10

-> X 4

-> X 25

-> X 28

-> X 22

-> X 19

-> X 16

-> X 13

-> X 3

-> X 7

X 1

-> X 20

-> X 17

-> X 4

-> X 25

-> X 28

-> X 22

-> X 19

-> X 16

-> X 13

-> X 3

X 1

-> X 6

-> X 14

-> X 17

-> X 20

-> X 22

-> X 19

-> X 16

-> X 4

-> X 25

-> X 28

start var = 1.51

X 1

-> X 3

-> X 6

-> X 25

-> X 22

-> X 19

-> X 17

-> X 20

-> X 23

-> X 24

-> X 9

X 1

-> X 4

-> X 3

-> X 6

-> X 25

-> X 22

-> X 19

-> X 17

-> X 20

-> X 23

X 1

-> X 2

-> X 5

-> X 3

-> X 6

-> X 25

-> X 22

-> X 19

-> X 17

-> X 20

-> X 23

start var = 2.01

X 1

-> X 5

-> X 8

-> X 14

-> X 11

-> X 22

-> X 25

-> X 28

-> X 17

-> X 20

-> X 23

X 1

-> X 15

-> X 12

-> X 6

-> X 9

-> X 25

-> X 28

-> X 17

-> X 20

-> X 23

-> X 29

X 1

-> X 3

-> X 17

-> X 20

-> X 23

-> X 29

-> X 26

-> X 8

-> X 14

-> X 11

-> X 22