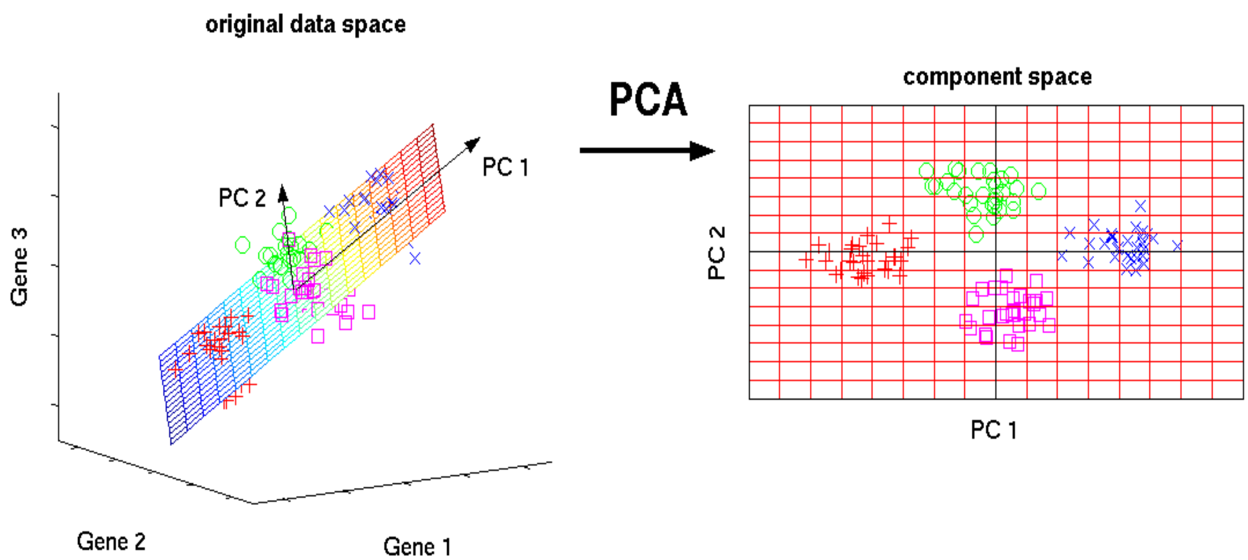
Logistic Regression Classification via the VIX, Part 2

By Philippe Rivet, Electrical Engineering '19, Western University

After introducing the VIX in 1993, the CBOE has over time added several variants based on investors' desire to modify their expectations of market volatility over different timescales. Notably, the 30-day period of the standard index is complemented by the VIX9D, VIX3M, VIX6M, and VIX1Y, with intervals of 9 days, 3 months, 6 months, and 1 year respectively. Within the one-month period the VVIX (implied forward-looking volatility of the VIX itself) is also used to make relevant decisions. As a continuation of the previous logistic regression effort, these indices are used to predict the occurrence of a 'recessionary event', i.e. a drop of 2.5% or more in the S&P/TSX composite index.

Principal Component Analysis (PCA)

In order to prevent the compute time of the machine learning model from becoming prohibitive, PCA can be leveraged to reduce the dimensions of the features used for parameter fitting. Basically this process involves minimizing the squared error from each dimensional component to yield an orthogonal basis which removes any redundant information while preserving uncorrelated data. Geometrically, this means establishing a local coordinate system where the axis are the perpendicular principal vectors. For this project the components are reduced from 5 to 2 for computation and visualization purposes.



Source: http://www.nlpca.org/fig_pca_principal_component_analysis.png

The steps for the process are as follows:

1. Calculate the covariance matrix of the data via

$$C = \frac{1}{m} X^T X$$

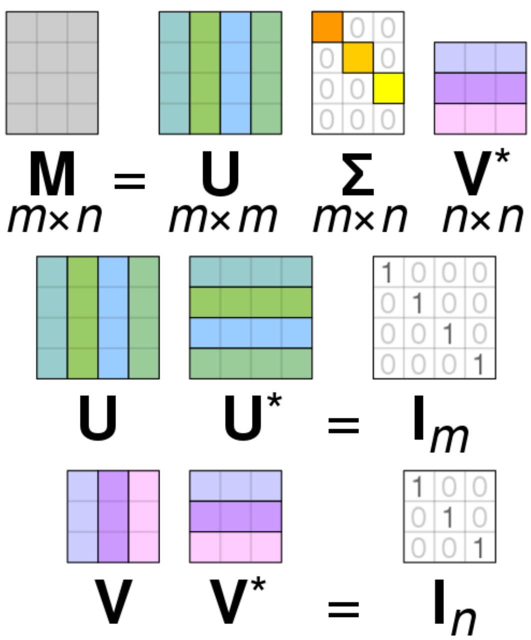
where m is the number of training examples and X is the data matrix (m x n).

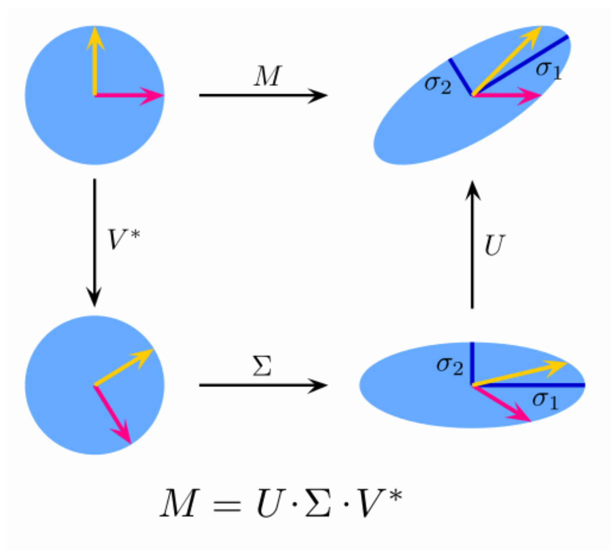
2. Perform Singular Value Decomposition, which is a method of factoring a matrix into a diagonal matrix containing its eigenvalues and two other rotation matrices.

$$S = U \Sigma V^T$$

Here U is an m x m orthogonal matrix, Σ is an m x n diagonalized matrix, and V is an n x n orthogonal matrix.

Illustration:





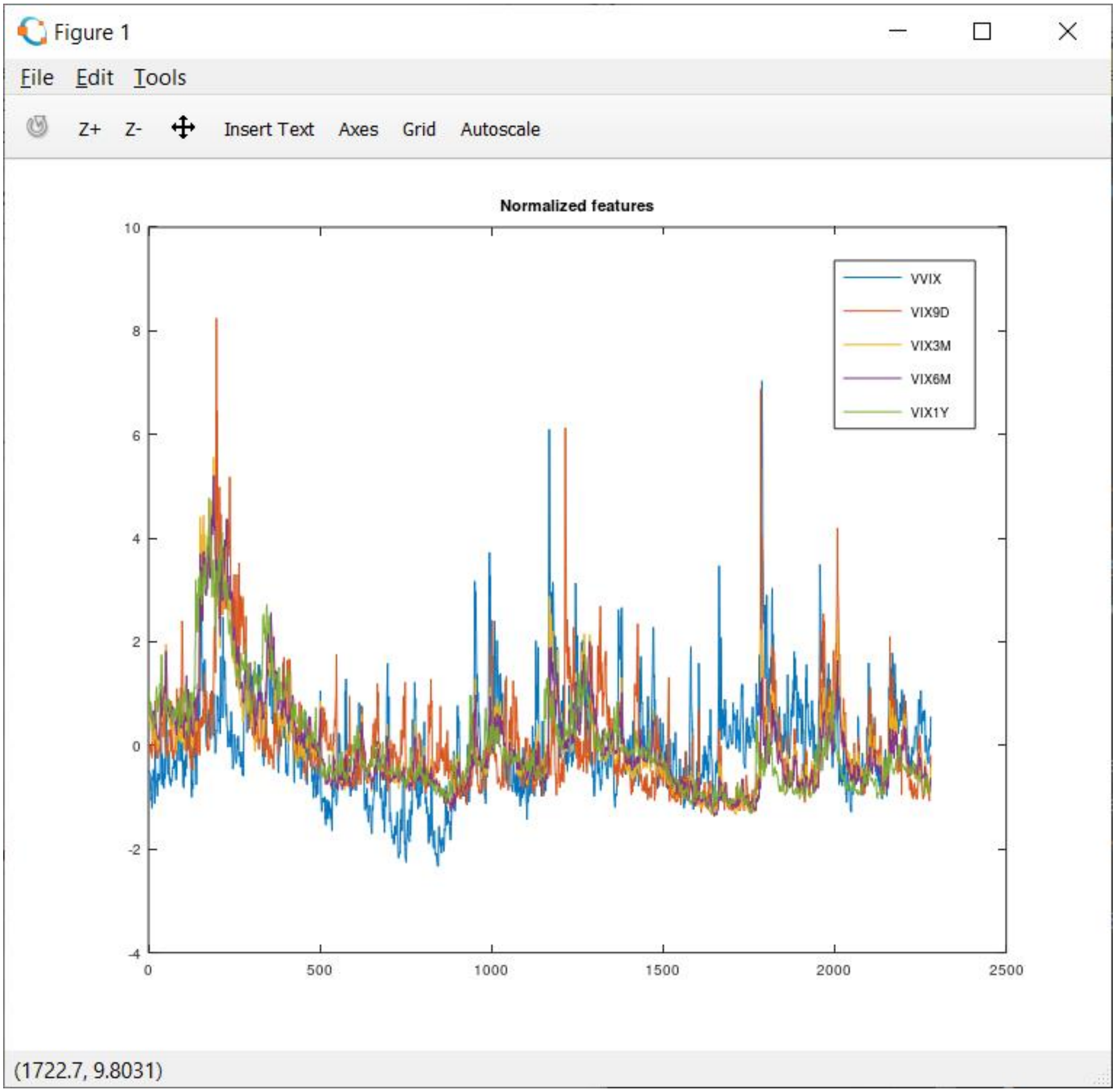
Sources:
https://upload.wikimedia.org/wikipedia/commons/thumb/c/c8/Singular_value_decomposition_visualisation.svg/ , https://upload.wikimedia.org/wikipedia/commons/thumb/b/bb/Singular-Value-Decomposition.svg/

Performing LR on Selected VIX variant data

Daily closing values of the VVIX, VIX9D, VIX 3M, VIX6M, VIX1Y, and S&P/TSX indices from January 3rd, 2011 to January 24th, 2020 were labelled and employed as training data for the binary classifier.

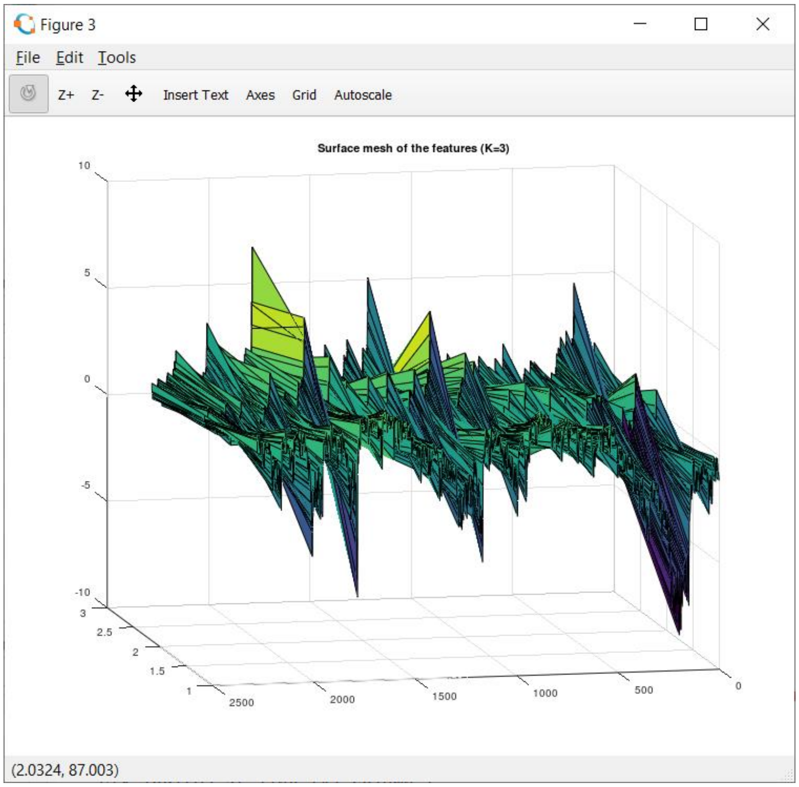
These features were 'normalized' as a pre-processing step, i.e. indivually scaled to fractions of the common mean for calculation efficiency later on.

An initial visualization produced the following plot:

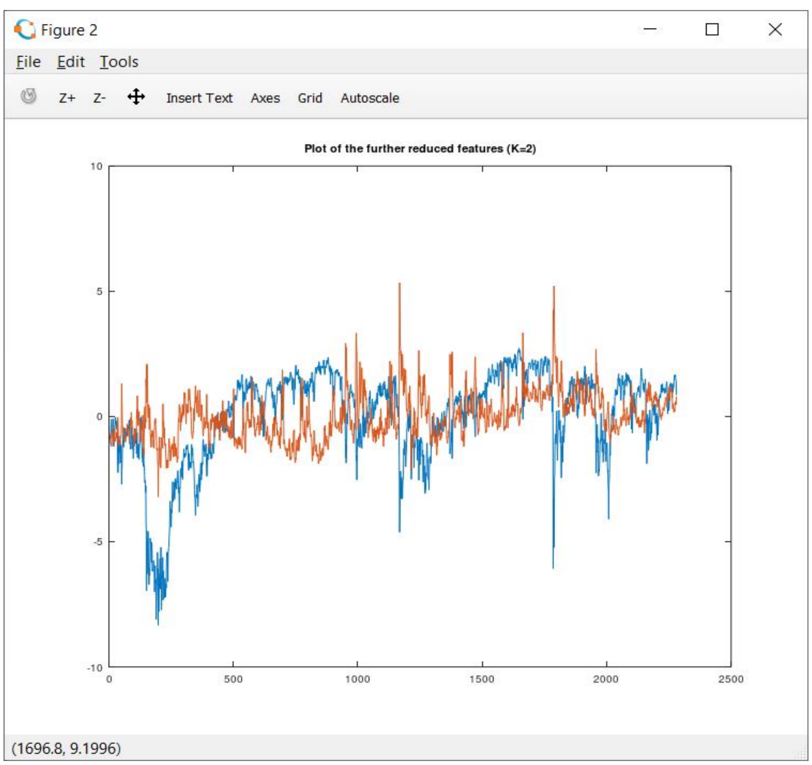


At several peaks these appear to be somewhat correlated, although there is significant variation in between.

Next, PCA was performed with the intent of simplifying the classification.



The above initial 3-dimensional plot indicated this was successful, so a further reduction to a single plane was executed:



Various runs of LR were performed with regularization, where the cost function is modified by adding an additional term lambda that penalizes overfitting:

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m} y^{(i)} \log\left(h_\theta\left(x^{(i)}\right)\right) + \left(1 - y^{(i)}\right) \log\left(1 - h_\theta\left(x^{(i)}\right)\right)\right] + \frac{\lambda}{2m}\sum_{j=1}^{n} \theta_j^2$$

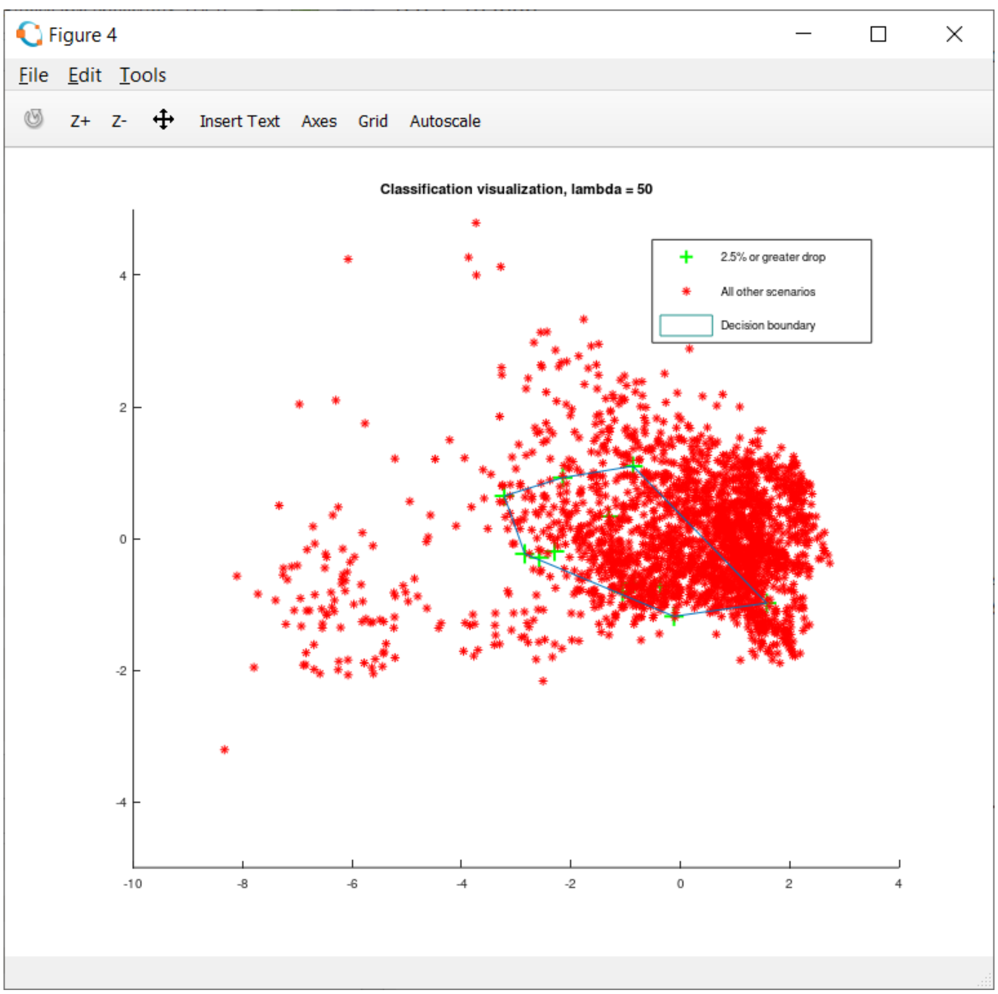
where λ is the regularization parameter that is tuned for the application in question.

The model parameter vector θ is as before updated through gradient descent:

$$\theta_{k2} = \theta_{k1} - \alpha\left[\sum_{i=1}^{m} (h_\theta\left(x^{(i)}\right) - y^{(i)})x_k^{(i)}\right] \quad (j = 0)$$

$$\theta_{k2} = \theta_{k1} - \alpha\left[\sum_{i=1}^{m} (h_\theta\left(x^{(i)}\right) - y^{(i)})x_k^{(i)}\right] + \frac{\lambda}{m}\theta_j \quad (j \geq 1)$$

After several iterations, a value of $\lambda = 50$ provided a satisfactory balance between a precise fit and generalization to future test data:



The clustering of the points in this plot potentially indicates that when they all experience a large change in magnitude a significant drop in the overall index is more likely. This prediction was validated when on March 16th the VVIX increased by 21.2%, the rest of the variants increased by approximately the same amount, and the S&P500 experienced the largest drop since 2008 of -11.98%.

Source: https://www.tradingview.com/chart/?symbol=CBOE%3AVVIX