

Machine Learning and Data Mining In Aerospace Technology Notes

Vivian Sedov

December 19, 2022

Contents

1	Tensor based Anomaly Detection	2
1.0.1	Tensor Classifier	3

1 Tensor based Anomaly Detection

Definition 1.1 Tensor Decomposition, also known as multi linear rank decomposition or tensor rank factorization. Is a mathematical technique for expressing a multi dimensional array tensor.

It is often used for dimensionality reduction, as it can reveal the underlying structure of the tensor, and identifying patterns and relationships that are not apparent when the tensor is expressed in its full form. A great example of this would be singular value decomposition, which is a generalization of the eigenvalue decomposition. In SVD, a tensor is decomposed into a sum of outer products of vectors, similar to the way that a matrix can be decomposed into a sum of outer products of its eigenvectors.

Here is an example of SVD on a matrix Say we have a matrix $m \times m$, and we want to find its svd

$$A = U \cdot S \cdot V^* * T$$

- Where U is a $m \times m$ Unitary matrix
- S is a $m \times n$ diagonal matrix
- V is the $n \times n$ unitary matrix

Further Explanation of the following quote: Which some what confused me

The target is to predict the observations labels into the test set. Thus, the developed model from the train set is utilized to predict the label (abnormal or normal) of observation into the test factor matrix. In [7] proposed to utilize the three-way data structure and apply a proper multi-way data analysis algorithm such as Parallel Factor Analysis, which is a simple model which obtained and utilized to train newness detectors. Such methods are evaluated both with simulated and real structural data to evaluate that the three-way analysis could be successfully utilized in structural health monitoring. Moreover, the advantage of such approach with regard to feature selection is also analyzed, Sensors make it possible to continually monitor pulses at multiple locations of a structure. Using a wide sensor network is useful for damage localization and a higher structural coverage, however it will also increase the number of variables. Thus, several dimensionality reduction is in demand, a PARAFAC decomposition accompanied by k number of components is utilized on the time- space- frequency tensor correlated to the normal samples and thereafter the acquired time factor matrix trained via k -NN (where features are the implicit variables). The model that has been built is then used for time point's classification in the incoming data.

Goal: In this context: the book states to use machine learning to predict the labels of observations in a test set that we get directly from a satellite. This is further based on patterns in a training set. The training set is consisted of a time series pattern, with multiple correlating variables, with corrected labels of the following :

- Normal
- Abnormal

For this to be achievable, it states to use a Tensor Decomposition method called *PARAFAC* to decompose the data tensor into a set of factors: and then using one of the factors - the time factor matrix : as an input to a machine learning algorithm like KNN. Which you should know is a supervised learning method. This will allow for classification of the labels.

PARAFAC, is being used to reduce the dimensionality of the data and extract the important patterns from it . And then using knn to learn a classification model based on those patterns.

Definition 1.2 – PARAFAC PARAFAC (Parallel Factors Analysis) is a tensor decomposition method that allows you to represent a multi-dimensional array (also known as a tensor) as the sum of simpler arrays. It is a generalization of the singular value decomposition (SVD), which is a widely used method for decomposing a matrix into a set of simpler matrices.

Like SVD, PARAFAC decomposes a tensor into a set of factors, which are lower-dimensional arrays that capture the most important patterns in the data. However, unlike SVD, which decomposes a matrix into a left singular matrix, a diagonal matrix, and a right singular matrix, PARAFAC decomposes a tensor into three or more factors, each of which corresponds to a different mode of the tensor.

For example, suppose you have a 3-dimensional tensor that represents a set of images, with the first mode representing the rows of the images, the second mode representing the columns of the images, and the third mode representing the color channels of the images. In this case, PARAFAC might decompose the tensor into three factors: one that represents the row structure of the images, one that represents the column structure of the images, and one that represents the color structure of the images.

PARAFAC is often used for tasks such as data compression, denoising, and data imputation, and can be especially useful for analyzing and understanding the structure of multi-dimensional data. It is also used in a variety of applications, including image processing, natural language processing, and structural health monitoring. With PARAFAC, we can decompose the matrix X into three factors A , B , and C such that:

$$X_{m,n} = \sum_{k=1}^r A_{m,k} \cdot B_{n,k} \cdot C_k$$

where r is the rank of the decomposition, and A , B , and C are matrices with shapes (m, r) , (n, r) , and $(r, 1)$, respectively.

With SVD, we can decompose the matrix X into a left singular matrix U , a diagonal matrix S , and a right singular matrix V such that:

$$X_{m,n} = U_{m,m} \cdot S_{m,n} \cdot V_{n,n}^T$$

where U and V are unitary matrices with shapes (m, m) and (n, n) , respectively, and S is a diagonal matrix with shape (m, n) .

So what you May want to think of parafac is getting the basis vector for a high level matrix. This is an extension of PCA, it is wise if you know what PCA is first before you move onto this.

1.0.1 Tensor Classifier

Definition 1.3 – Tensor classifiers Tensor classifiers are machine learning models that are specifically designed to handle tensorial (multi-dimensional array) data. These models can be used to classify data points based on their features, which are represented as tensors.

One example of a tensor classifier is the support vector machine (SVM), which has been

extended to handle tensorial data in the form of a tensor support vector machine (TSVM). TSVMs are trained directly on tensorial data, and the resulting model can be used for prediction. Tensor classifiers are particularly useful for detecting anomalies in multiway data, and have been applied successfully to image classification tasks.

Another example of a tensor classifier is the supervised tensor learning (STL) framework, which is a combination of multilinear algebra operations and convex optimization. STL can be used to generalize a variety of classic machine learning techniques, such as support vector machines, Fisher discriminant analysis, and distance metric learning, to handle tensorial data.

Tensor classifiers differ from traditional vector-based classifiers in that they represent training measurements as tensors rather than vectors, and the classification decision function is defined differently. In vector-based learning, the classification decision function is defined by a hyperplane, while in tensor-based learning, it is defined by a tensorplane. Tensor classifiers may be particularly useful for small sample cases, as they require fewer parameters to be estimated than traditional vector classifiers.

Vector-based learning: In vector-based learning, the training measurements are represented as vectors, and the classification decision function is defined by a hyperplane. The hyperplane is defined by a weight vector w and a bias term b , such that the classification function is given by: $y(x) = \text{sign}[w^T x + b]$ where x is a feature vector and y is the predicted class label. Tensor-based learning:

In tensor-based learning, the training measurements are represented as tensors, and the classification decision function is defined by a tensorplane. The tensorplane is defined by weight tensors w_k (for $1 \leq k \leq M$) and a bias term b , such that the classification function is given by: $y(X) = \text{sign}\left[\sum_{k=1}^M X_k \cdot w_k + b\right]$ The Tensor Least Square (TLS) method is an extension of the least squares classifier, which is a type of linear regression model. TLS is a tensor-based classifier that estimates the classification function by minimizing the sum of squared errors between the predicted and actual class labels. TLS has been tested on six databases from the UCI repository, and has been shown to be particularly suitable for small sample cases due to its lower number of required parameters compared to traditional vector classifiers.