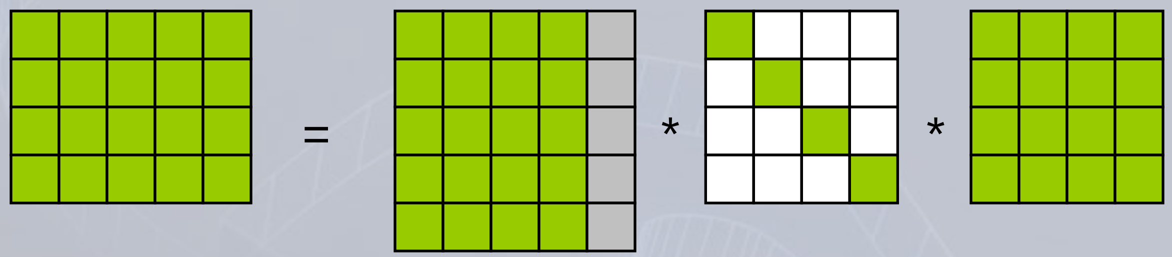
Singular Value Decomposition (SVD)

SVD of a Matrix



$$M = U \quad S \quad V^T$$

U and V are orthogonal matrices, and S is a diagonal matrix consisting of singular values.

Singular Value Decomposition (SVD)

SVD of a Matrix: observations

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

Multiply both sides by M^T

Multiplying on the left

$$M^TM = (U\ S\ V^T)^T\ U\ S\ V^T$$

$$M^TM = (V\ S\ U^T)\ U\ S\ V^T$$

$$U^TU = I$$

$$M^TM = V\ S^2\ V^T$$

Multiplying on the right

$$MM^T = U\ S\ V^T\ (U\ S\ V^T)^T$$

$$MM^T = U\ S\ V^T\ (V\ S\ U^T)$$

$$V^TV = I$$

$$MM^T = U\ S^2\ U^T$$

Singular Value Decomposition (SVD)

SVD of a Matrix: observations

$M^TM = V\ S^2\ V^T$ ⟵ diagonalizations ⟶ $MM^T = U\ S^2\ U^T$

Diagonalization of a Matrix: (finding eigenvalues)

$A = W\ \Lambda\ W^T$
where:
- A is a square, symmetric matrix
- Columns of W are eigenvectors of A
- Λ is a diagonal matrix containing the eigenvalues

Therefore, if we know U (or V) and S, we basically have found out the eigenvectors and eigenvalues of MM^T (or M^TM) !

Principal Component Analysis (PCA)

What is PCA?

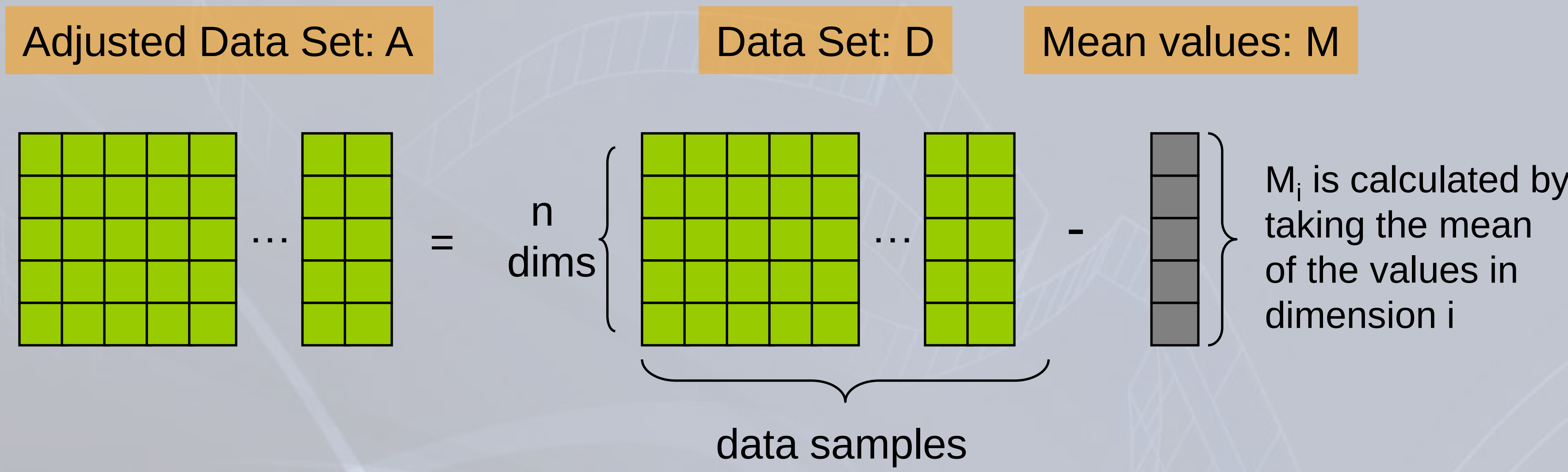
- Analysis of n-dimensional data
- Observes correspondence between different dimensions
- Determines principal dimensions along which the variance of the data is high

Why PCA?

- Determines a (lower dimensional) basis to represent the data
- Useful compression mechanism
- Useful for decreasing dimensionality of high dimensional data

Principal Component Analysis (PCA)

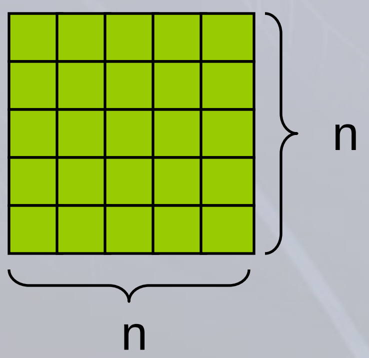
Steps in PCA: #1 Calculate Adjusted Data Set



Principal Component Analysis (PCA)

Steps in PCA: #2 Calculate Co-variance matrix, C, from Adjusted Data Set, A

Co-variance Matrix: C



$C_{ij} = cov(i,j)$

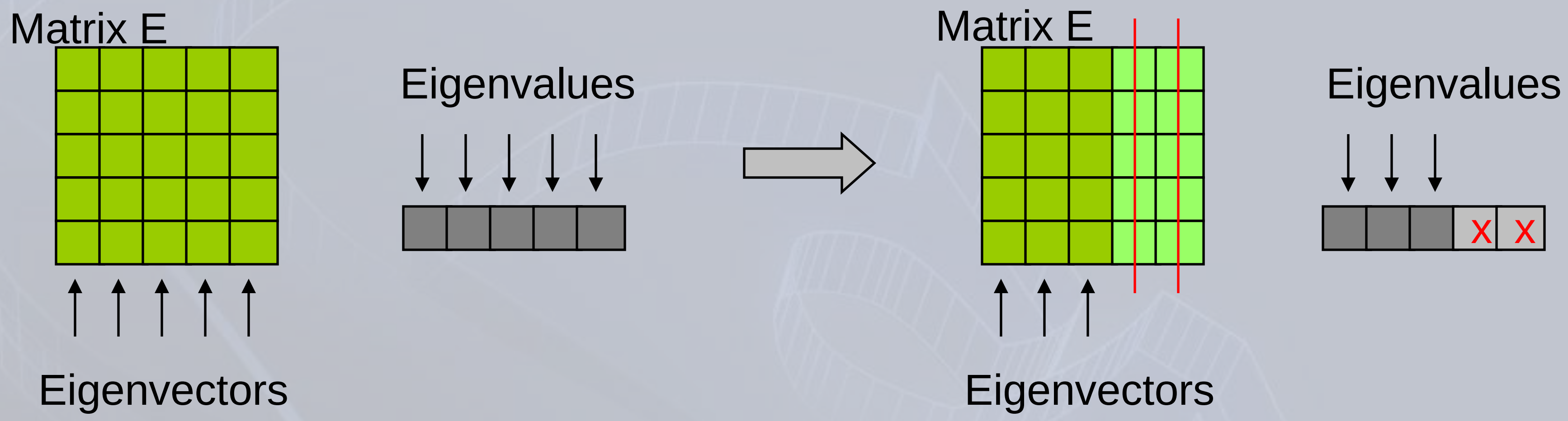
$$cov(X,Y) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)}$$

Note: Since the means of the dimensions in the adjusted data set, A, are 0, the covariance matrix can simply be written as:

$$C = (A\,A^T) / (n-1)$$

Principal Component Analysis (PCA)

Steps in PCA: #3 Calculate eigenvectors and eigenvalues of C



If some eigenvalues are 0 or very small, we can essentially discard those eigenvalues and the corresponding eigenvectors, hence reducing the dimensionality of the new basis.

Principal Component Analysis (PCA)

Steps in PCA: #4 Transforming data set to the new basis

$F = E^T A$

where:

- F is the transformed data set
- E^T is the transpose of the E matrix containing the eigenvectors
- A is the adjusted data set

Note that the dimensions of the new dataset, F, are less than the data set A

To recover A from F:

$(E^T)^{-1}F = (E^T)^{-1}E^T A$

$(E^T)^T F = A$

$EF = A$

* E is orthogonal, therefore $E^{-1} = E^T$

PCA using SVD

Recall: In PCA we basically try to find eigenvalues and eigenvectors of the covariance matrix, C. We showed that $C = (AA^T) / (n-1)$, and thus finding the eigenvalues and eigenvectors of C is the same as finding the eigenvalues and eigenvectors of AA^T

Recall: In SVD, we decomposed a matrix A as follows:

$$A = U S V^T$$

and we showed that:

$$AA^T = U S^2 U^T$$

where the columns of U contain the eigenvectors of AA^T and the eigenvalues of AA^T are the squares of the singular values in S

Thus SVD gives us the eigenvectors and eigenvalues that we need for PCA