# Principal Component analysis

Lukas Käll
lukask@kth.se

# Decomposition of a matrix

- From linear algebra, we know that we can multiply two vectors into a matrix. $\mathbf{M} = \mathbf{u} \otimes \mathbf{v} = \mathbf{u}\mathbf{v}^\top = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & u_1 v_3 \\ u_2 v_1 & u_2 v_2 & u_2 v_3 \\ u_3 v_1 & u_3 v_2 & u_3 v_3 \\ u_4 v_1 & u_4 v_2 & u_4 v_3 \end{bmatrix}$.

# Decomposition of a matrix

- From linear algebra, we know that we can multiply two vectors into a matrix. $\mathbf{M} = \mathbf{u} \otimes \mathbf{v} = \mathbf{u}\mathbf{v}^{\mathsf{T}} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & u_1 v_3 \\ u_2 v_1 & u_2 v_2 & u_2 v_3 \\ u_3 v_1 & u_3 v_2 & u_3 v_3 \\ u_4 v_1 & u_4 v_2 & u_4 v_3 \end{bmatrix}$.
- What if we could do the opposite? I.e. given a matrix $\mathbf{M}$, what would be the vectors $\mathbf{u}$ and $\mathbf{v}$ that closest represents $\mathbf{M}$ so that $\mathbf{M} \approx \mathbf{u}\mathbf{v}^{\mathsf{T}}$. This is in essence what you do with principal component analysis (PCA).

# Conspicuous Example

- See notebook

# More principal components to your PCA

- Once you remove the principal components from a matrix $\mathbf{M}$ the remaining residues,i.e. $\mathbf{M} - \mathbf{u}^{(1)}\mathbf{v}^{(1)\mathbf{T}}$ might in its turn be decomposed into vectors. I.e. we can calculate the vectors $\mathbf{u}^{(2)}$ and $\mathbf{v}^{(2)}$ that best describe the matrix $\mathbf{M} - \mathbf{u}^{(1)}\mathbf{v}^{(1)\mathbf{T}}$. We call these the second principal components, and the first are called first principal components.

# More principal components to your PCA

- Once you remove the principal components from a matrix $\mathbf{M}$ the remaining residues,i.e. $\mathbf{M} - \mathbf{u}^{(1)}\mathbf{v}^{(1)\mathbf{T}}$ might in its turn be decomposed into vectors. I.e. we can calculate the vectors $\mathbf{u}^{(2)}$ and $\mathbf{v}^{(2)}$ that best describe the matrix $\mathbf{M} - \mathbf{u}^{(1)}\mathbf{v}^{(1)\mathbf{T}}$. We call these the second principal components, and the first are called first principal components.
- In this manner we can derive as many principal components as there are rows or columns (which ever is the lowest number) in $\mathbf{M}$. In most applications, we settle for two such components.

# Thanks!