Week 8: take-home messages

- Decide the optimal number of principal components: those that explain a desired proportion of the total variance.
- When there is no correlation between features, the covariance matrix closely resembles a diagonal matrix. This means that small and comparable eigenvalues are expected, and PCA will not have much impact.
- Plotting data points in 2D after applying PCA may reveal patterns and relationships between classes.

- Update rule: first update H, then update W with the new H.
- When dividing by some quantity that can be zero, we can add a small constant to prevent numerical errors.
- Understand different matrix multiplication operations (e.g. `@`, `*`, `numpy.dot`, `numpy.matmul`).