Support vector machines unit:

- What PCA is (Monday)
- Applications of PCA (Wednesday, in lab)
- The math of PCA (**Today**)
- PCA algorithms (next week Monday)

Today:

- Follow-up from yesterday's lab
- Maximum-variance view
  - Finding the first (or zeroth) principal component
  - Finding more principal components
- Minimum-information-loss view

The most important source for all of this was Deisenroth et al, *Mathematics for Machine Learning*, 2020, 286–293.

(Available on Canvas)

Coming up:

Due Fri, Mar 28:
Textbook reading from Chapter 10 (see Canvas)

Due Mon, Mar 31:
Take PCA quiz

Due Fri, Apr 4:
Implement PCA

Sometime between Mar 31 and Apr 17:
Make an office-hours appointment for project check-in
(Originally the deadline was Apr 11)

(There will be an ethics reading thrown in here sometime)

Let $\mathbf{W} = \begin{pmatrix} \boldsymbol{v_0} \\ \vdots \\ \boldsymbol{v_{M-1}} \end{pmatrix} \in \mathbb{R}^{M \times D}$ be a set of orthonormal basis vectors.

Let $\boldsymbol{z_n} = \mathbf{W}\boldsymbol{x_n}$ be the projection of $\boldsymbol{x_n}$ in the vector space defined by these basis vectors.

Let $\mathbf{U} = \mathbf{W}^T$. We can project $\boldsymbol{z_n}$ back ito the original space with $\boldsymbol{\tilde{x}_n} = \mathbf{U}\boldsymbol{z_n}$.

Let $\boldsymbol{v}_0 \in \mathbb{R}^D$ be the zeroth ("first") principal component (which we want to find).

Let $z_{0,n} = \boldsymbol{v_0}^T \boldsymbol{x_n}$ be the zeroth coordinate in the projection of $\boldsymbol{x_n}$.

Let z_0 be a random variable modeling the value of the zeroth coordinate. Assume data is centered, that is, $\mathbb{E}[z_0] = 0$. Then

$$
\begin{aligned}
\mathrm{Var}[z_0] &= \tfrac{1}{N} \textstyle\sum_{n=0}^{N-1} (z_{0,n} - 0)^2 \\
&= \tfrac{1}{N} \textstyle\sum_{n=0}^{N-1} \left(\boldsymbol{v_0}^T \boldsymbol{x_n}\right)^2 \\
&= \tfrac{1}{N} \textstyle\sum_{n=0}^{N-1} \left(\boldsymbol{v_0}^T \boldsymbol{x_n} \boldsymbol{x_0}^T \boldsymbol{v_n}\right) \\
&= \boldsymbol{v_0}^T \left(\tfrac{1}{N} \textstyle\sum_{n=0}^{N-1} \boldsymbol{x_n} \boldsymbol{x_n}^T \right) \boldsymbol{v_0}
\end{aligned}
$$

Let $\boldsymbol{v}_0 \in \mathbb{R}^D$ be the zeroth ("first") principal component (which we want to find).

Let $z_{0,n} = \boldsymbol{v_0}^T \boldsymbol{x_n}$ be the zeroth coordinate in the projection of $\boldsymbol{x_n}$.

Let z_0 be a random variable modeling the value of the zeroth coordinate. Assume data is centered, that is, $\mathbb{E}[z_0] = 0$.

Let $\mathbf{C} = \frac{1}{N} \sum_{n=0}^{N-1} \boldsymbol{x_n} \boldsymbol{x_n}^T$ be the data covariance matrix. Then

$$
\begin{aligned}
\text{Var}[z_0] &= \frac{1}{N} \sum_{n=0}^{N-1} (z_{0,n} - 0)^2 \\
&= \frac{1}{N} \sum_{n=0}^{N-1} (\boldsymbol{v_0}^T \boldsymbol{x_n})^2 \\
&= \frac{1}{N} \sum_{n=0}^{N-1} (\boldsymbol{v_0}^T \boldsymbol{x_n} \boldsymbol{x_0}^T \boldsymbol{v_n})^2 \\
&= \boldsymbol{v_0}^T \left(\frac{1}{N} \sum_{n=0}^{N-1} \boldsymbol{x_n} \boldsymbol{x_n}^T \right) \boldsymbol{v_0} \\
&= \boldsymbol{v_0}^T \mathbf{C} \boldsymbol{v_0} \qquad \text{which we want to maximize subject to } ||\boldsymbol{v_0}||^2 = 1
\end{aligned}
$$

Optimization problem as a Lagrangian:

$$\mathcal{L}(\boldsymbol{v_0}, \lambda_0) = \boldsymbol{v_0}\mathbf{C}\boldsymbol{v_0} + \lambda_0(1 - \boldsymbol{v_0}^T\boldsymbol{v_0})$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_0} = 1 - \boldsymbol{v_0}^T\boldsymbol{v_0}$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \boldsymbol{v_0}} &= 2\boldsymbol{v_0}^T C - 2\lambda_0\boldsymbol{v_0}^T &= 0 \\ \boldsymbol{v_0}^T C &= \lambda_0\boldsymbol{v_0}^T \end{aligned}$$

Plug that into our formula for the variance:

$$\begin{aligned} \mathrm{Var}[z_0] &= \boldsymbol{v_0}^T\mathbf{C}\boldsymbol{v_0} \\ &= \lambda_0\boldsymbol{v_0}^T\boldsymbol{v_0} &= \lambda_0 \end{aligned}$$

Coming up:

Due Fri, Mar 28:
Textbook reading from Chapter 10 (see Canvas)

Due Mon, Mar 31:
Take PCA quiz

Due Fri, Apr 4:
Implement PCA

Sometime between Mar 31 and Apr 17:
Make an office-hours appointment for project check-in
(Originally the deadline was Apr 11)

(There will be an ethics reading thrown in here sometime)