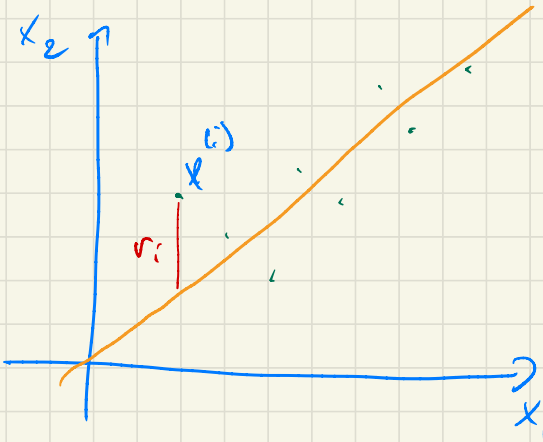
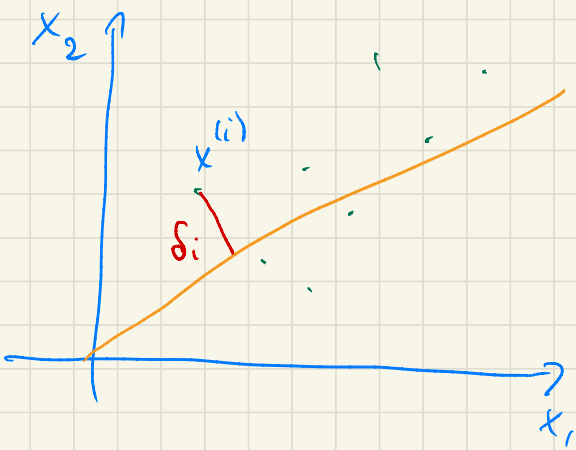
8.4.2020 PCA



In linear regression, minimize

$$\sum_i \underbrace{\left(x_2^{(i)} - w \cdot x_1^{(i)}\right)^2}_{r_i}$$

$w \in \mathbb{R}$



In PCA, minimize

$$\sum_i \underbrace{\left\| x^{(i)} - w \cdot (w^T x^{(i)}) \right\|^2}_{\delta_i}$$

$w \in \mathbb{R}^2$

Empirical covariance of $D = \{x_1, \dots x_n\}$, $x_i \in \mathbb{R}^d$ is $\Sigma = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu})(x_i - \hat{\mu})^T$, where $\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i$. In PCA, assume data is centered (i.e., $\hat{\mu} = 0$)