Lecture 10 Outline:

1. Review: Logistic Regression.

2. Multinomial Regression (Multi-class logistic regression).
   - $K - 1$ formulation: $p_k = \frac{\exp(X\beta_k)}{(1 + \sum_{l=1}^{K-1} \exp(X\beta_l))}$
   - $K$ formulation: $p_k = \frac{\exp(X\beta_k)}{\sum_{l=1}^{K} \exp(X\beta_l)}$.

   - Penalized log-likelihood.
   - Penalized multinomial log-likelihood & group lasso penalty.
   - Fitting Sparse Logistic Models.
     - Proximal Newton methods.

4. Optimal Separating Hyperplanes (Maximum Margin Classifiers).
   - Separating hyperplanes for linearly separable classes.
   - Defining hyperplanes, margins, and +/- planes.
   - Primal form optimization problem.
   - Interpretation: support vectors.
   - Behavior in practice.