Machine Learning (CS 181):
Style Guide
Basic Notation

\[ A \in \mathbb{R}^{m \times n} \] matrices are bold caps
\[ a \in \mathbb{R}^{m(\times 1)} \] vectors are bold lower, always column
\[ a^\top \] transpose
\[ a \in \mathbb{R} \] scalars are lower case, non bold
\[ \mathcal{A} \] sets are script case
\[ \{ a_i \}_{1}^{n} \] a sequence of \( a_1 \ldots a_n \)

- We distinguish between \( a_i \) and \( b_i \). The first is the \( i^\prime \text{th} \) vector of a sequence, the second is the \( i^\prime \text{th} \) scalar in \( b \).
Supervised Learning

\[ \mathbf{X} \in \mathbb{R}^{n \times m} \]  \( n \) training instances with \( m \) features

\[ X_{i,j} \]  the \( j \)'th feature of example \( i \)

\[ \mathbf{y} \in \mathbb{R}^{n} \]  the \( n \) target instances (regression)

\[ \mathbf{y} \in \{-1, 1\}^{n} \]  the \( n \) target instances (classification)

- When describing data as a sequence we use

\[ D = (x_1, y_1) \ldots (x_n, y_n) = \{(x_i, y_i)\}_1^n \]
Models

\[ w \in \mathbb{R}^m \quad \text{linear-model parameters} \]
\[ f(x; w) \quad \text{model parameterized by } w \]
\[ \hat{y} \quad \text{model prediction (to be distinguished from target } y) \]
\[ \phi : \mathbb{R}^m \leftrightarrow \mathbb{R}^d \quad \text{basis function (if used changes } w \text{ dim) } \]
Optimization

- $\mathcal{L}(w)$; loss function
- $w^* = \arg \min_w \mathcal{L}(w)$; training minimization
- $\mathcal{L}(w) = -\sum_{i=1}^{n} \ln p(y_i | x_i; w)$; negative log-likelihood
- $\min_x x \quad \text{s.t.} \quad Ax = b$; constrained
Probabilistic Notation

- $p(y | x)$; discrete distribution $p(Y = y | X = x)$ (RVs implied)
- $p(y | x) \propto p(x | y)p(y)$; proportional to notation
- $\mathcal{N}(x | \mu, \Sigma)$; multivariate normal (prefer $\exp$ to $e$)
  \[
  \mathcal{N}(x | \mu, \Sigma) = \frac{1}{\sqrt{|2\pi\Sigma|}} \exp\left(-\frac{1}{2}(x - \mu)\top\Sigma^{-1}(x - \mu)\right)
  \]
- $\alpha, \beta, \gamma$; scalar hyperparameters (explicitly conditioned $p(w|\alpha)$)
- $\alpha, \beta, \gamma$; vector hyperparameters