Multivariate Regression for Time Series

Sept 4, 2019

- Definitions
  - $x_i \in \mathcal{X} \subseteq \mathbb{R}^p$
  - $y_i \in \mathcal{Y} \subseteq \mathbb{R}^q$
  - Given data $\mathcal{D}_n = (x_i, y_i)_{i=1}^n$
  - Assume $(x_i, y_i) \overset{iid}{\sim} F_{XY}$
  - Learn $F_{Y|X}$ from $\mathcal{D}_n$

- Univariate regression:
  - Given $y \in \mathbb{R}$, estimate $\mathbb{E}(Y|X)$

- Multivariate regression:
  - Given $y \in \mathbb{R}^q$, estimate $\mathbb{E}(Y|X)$

- Multivariate time series regression (forecasting)
  - Assume $n=1$:

    $X \in \mathbb{R}^{p \times t}$ (p sensors, t timepoints)

    $y \in \mathbb{R}^{1 \times t'}$ or $y \in \mathbb{R}^{p \times t'}$ (collapse all channels into a single output stream)
  - For $n > 1$, $x \in \mathbb{R}^{p \times t_i}$
  - Goal:
    - Learn $h(\cdot|\mathcal{D}_n) : \mathbb{R}^{p \times t} \to \mathcal{Y}$ st h minimizes $\mathbb{E}_{x,y}[(Y - h(X|\mathcal{D}_n))^2]$