Lecture 6: Least Squares and RLS

Zach Neveu

October 10, 2019

Course Structure Notes

- DO ALL EXERCISES
- Exercises are primary material for exam - this is why no solns given
- Pset 2: due 27/10 - supposed to be pretty easy
- Mini Project - max 3 pages, 1-3 people, presented at oral exam
  - Focus on using a different dataset, and adding some extra analysis
  - 15-20 hours per person
  - One week at the end with no lecture for focusing on this
- Exam sometime after 12/10

Review

- Orthogonality principle: $\hat{y}$ vector should be orthogonal to the plane of the data points
- In other words, $e * \hat{y} = 0$ because $\hat{y}$ is orthogonal to e
- NLMS algorithm. Theories around valid, stable and optimal step sizes μ
  - Step size: between $0 < \mu < \frac{1}{\lambda_{max}}$, λ is max eigval of Σ_x
  - Stable step size: $\mu < \frac{2}{tr(\Sigma_x)}$ (trace is sum of eigvals)
  - Total error: $J_{total} = J_{min} + J_{excess} + J_{lag}$
  - Excess error: error resulting from the fact that SGD never fully stabilizes (variance of converged algorithm)
  - Min error: error once algorithm is stable. (bias in optimal point)
  - Lag error: In time-varying environment, filter always lags behind reality, causing error. Basically, the optimal point keeps changing.
  - Big step size adds to excess error, subtracts from lag

RLS

- Converges in less iterations than NLMS
- Has better error rate than NLMS

- How does it get this free lunch??
- Loses time tracking ability, not as good in real life sometimes.
- Possible to diverge if large changes occur
- RLS quite sensitive to fixed point errors - big problem in hardware
- Cost function exponentially weights
- Params:
  - $\beta \rightarrow$ forgetting factor. Small $\beta \rightarrow$ forget fast
  - $\lambda \rightarrow$ regularization multiplier - when few n, don't overfit

$\rightarrow$