**Reference:**

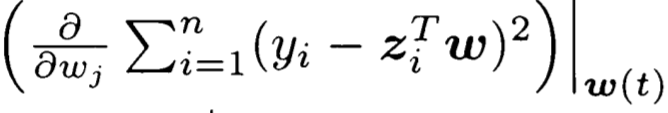
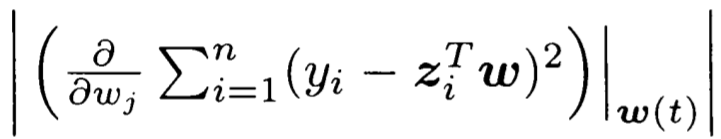
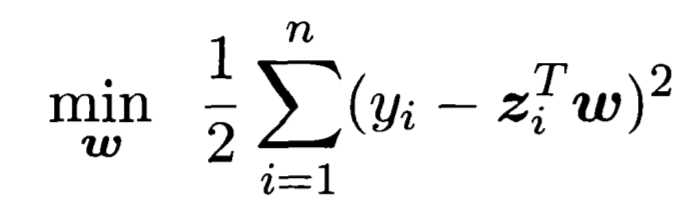
Wu (2012), Elastic Net for Cox’s Proportional Hazards Model with a Solution Path Algorithm.

**Summary:**

* LAR (least angle regression) is a solution path algorithm for least squares regression.
* Slight modification of LAR leads to whole LASSO solution path.
* LAR and LASSO solution paths are piecewise linear.
* We extend the LAR to handle Cox’s PH model.
* Goal: develop a solution path algorithm for the elastic net penalty in Cox’s PH model.
  + This is done by:

1. Extending LAR to optimize the log partial likelihood + a fixed small ridge term.
2. Defining a path modification which leads to the solution path of the elastic net regularized log partial likelihood.
   * The solution path is exact and piecewise determined by ordinary differential equation systems.

**Introduction:**

* Goals of survival analysis:
  + Characterize dependence of survival time Y on X predictors.
  + Identify important risk factors and quantify their risk contributions.
    - Solution: model selection (e.g. best-subset selection, stepwise selection, LASSO, SCAD, etc.) (select a model with a certain number of variables)
    - LASSO (L1) penalty: leads to biased estimates for true non-zero coefficients.
    - SCAD penalty: is symmetric and piecewise quadratic, and solves the L1 biased issue.
* LAR (least angle regression) solution path algorithm: produces a piecewise linear solution path for the least squares regression.
* LARS: LASSO and Forward Stagewise linear regression solution paths, slight modifications of the LAR
  + For OLS (ordinary least squares), the intermediate solution path is piecewise linear; over each piece, it moves along the direction that keeps the correlation between the current residuals and each active predictor equal in absolute value.
  + This implies that the objective function has the same absolute value of the first-order partial derivatives for each active predictor along the LAR solution path.
* Figure 1:
  + Top-left panel: the LAR path w\_j(t) against the relative one-norm |w(t)|/|max(w(t))|, for each predictor “j” (“w” are the “beta” coefficients)
    - The relative one-norm is each “LAR step”
  + Top-right panel: the first-order partial derivatives of of each predictor.
  + Bottom panel: the absolute value of the first-order partial derivatives of each predictor.
  + At the end of each LAR step, a new predictor “w” joins the group of active predictors, sharing the honor of having the same largest absolute value of the first-order partial derivatives.
  + The LAR algorithm ends at the full OLS estimate of  when all first-order partial derivatives are exactly zero.
* We extend the LAR algorithm to Cox’s PH model.
  + CoxLAR-ridge: adding a fixed small ridge term to the log partial likelihood function.
  + CoxLAR: when the ridge term is zero, you have only the original log partial likelihood function.
  + CoxEN: the LASSO penalty with a small ridge term.
  + CoxLASSO: setting the ridge term to be zero in elastic net.
* Advantages of elastic net:
  + Capable of selecting more predictors than the sample size (p > n)
  + In LASSO, # of predictors selected can at most be equal to sample size (p = n)
* Park and Hastie’s solution path algorithm for L1 GLMs and Cox’s PH model:
  + Algorithm based on the predictor-corrector method of convex optimization.
  + Cons:
    - They still need to solve many optimization problems, one at each tuning parameter point. (1)
    - They did not address how the solution changes when the tuning parameter changes. (2)
* Solution: new algorithms CoxLAR and CoxLASSO.
  + Addresses con (2), the solution path propagates according to ordinary differential equation (ODE) systems.
  + The commonly used fourth-order Runge-Kutta method can be used to solve these ODE systems to obtain the whole CoxLARS solution paths.

LARS Algorithm (for linear regression models):

* Set all Beta coefficients to 0.
* Find the predictor x\_j that is most correlated with y (or “r” residuals).
* Increase B\_j in the direction of the sign (positive or negative) of it’s correlation with y. Take residuals along the way. (e.g. B\_j = 1, 2, 3…)
* Stop increasing B\_j when some other predictor x\_k has as much correlation with “r” as x\_j has.
* Extend (B\_j, B\_k) in a direction that is equiangular to both x\_j and x\_k.
* Repeat until all predictors are in the model.

**Numerical Examples:**

Example 1

* Use a simulated dataset to demonstrate that true LASSO regularized solution path is NOT piecewise linear.
* **X** is N(0, sigma), sigma is 3x40.
* h(y|x) “lifetime for an individual” are generated conditional on **X** for an individual, **B**, and h\_0(y).
* Censoring time was uniformly distributed U[0,8], censoring rate was 32.3%.
* Figure 2:
  + Top-left panel:
    - Solid lines show the CoxLASSO solution path, B(t) with respect to the one-norm |B(t)|.
    - Dashed straight lines show what a solution path would be like if it were piecewise linear.
  + Top-right panel:
    - First-order partial derivatives along CoxLASSO solution path (first-order with respect to one-norm).
  + Bottom-left panel:
    - Absolute value of first-order partial derivatives with respect to one-norm.
  + Bottom-right panel:
    - Absolute value of first-order partial derivatives with respect to t.
* What is t?

Example 2

* Demonstrates how the Cox-LASSO modification leads to the CoxLASSO path when ridge term is 0, and how it leads to CoxEN when ridge term is > 0.
* Generated **X**, h\_0(y), **B**.
* Censoring time generated from U[0,10], leading to a censoring rate of 30.5%.
* Figure 3:
  + Top-left panel:
    - Solution path of CoxLAR when ridge term is 0.
  + Bottom-left panel:
    - Solution path of CoxLASSO when ridge term is 0.
  + Top-right and bottom-right panels:
    - Same thing but with ridge term = 0.2.

Example 3

* Studies the dependency of survival time on 17 covariates.
* Standardized each covariate to have mean 0 and variance 1.
* CoxLARS was applied to standardized data with all 17 variables.
* With ridge term = 0, CoxLAR and CoxLASSO gave the same solution path.