Fast, Exact Model Selection and Permutation Testing for ℓ_2-Regularized Logistic Regression

Bryan Conroy
Columbia University
New York, NY
bc2468@columbia.edu

Paul Sajda
Columbia University
New York, NY
psajda@columbia.edu

1 APPENDIX - SUPPLEMENTARY MATERIAL

Proof. (Lemma 1): Given the linear system $A_p w_p = X m_p$, with $A_p = X R_p X^T + C$ and $X = QZ$. Let the columns of $Q^{\perp}$ span the orthogonal complement of range(Q). Then we have:

$$\begin{aligned} {Q^{\perp}}^T (X R_p X^T + C) w_p &= {Q^{\perp}}^T X m_p \quad (1) \\ {Q^{\perp}}^T C w_p &= 0 \quad (2) \end{aligned}$$

This implies that $C w_p \in$ range(Q), so that $w_p \in$ range($C^{-1}Q$). □

Proof. (Lemma 2): Let $U = [Q, W]$ be a matrix with orthonormal columns that is a basis for range($[Q, C^{-1}Q]$). From Lemma 1, we know that the solution w_p to $A w_p = X m_p$ satisfies $w_p \in$ range(U). Thus, write w_p as:

$$w_p = Q\tilde{w}_p + W\alpha_p \quad (3)$$

Then we may factorize $A_p w_p = X m_p$ as:

$$\begin{aligned} U^T (X R_p X^T + C) U \begin{bmatrix} \tilde{w}_p \\ \alpha_p \end{bmatrix} &= U^T X m_p \\ \begin{bmatrix} Z R_p Z^T + Q^T C Q & Q^T C W \\ W^T C Q & W^T C W \end{bmatrix} \begin{bmatrix} \tilde{w}_p \\ \alpha_p \end{bmatrix} &= \begin{bmatrix} Z \\ 0 \end{bmatrix} m_p \end{aligned}$$

The above may be separated into two systems of linear equations:

$$\begin{aligned} (Z R_p Z^T + Q^T C Q)\tilde{w}_p + Q^T C W \alpha_p &= Z m_p \quad (4) \\ W^T C Q \tilde{w}_p + W^T C W \alpha_p &= 0 \quad (5) \end{aligned}$$

Using the second system of equations to solve for α_p in terms of $\tilde{w}_p$ yields:

$$\alpha_p = -F\tilde{w}_p \quad (6)$$

where $F = -(W^T C W)^{-1} W^T C Q$.

Substituting (6) into (4) yields:

$$\tilde{A}_p \tilde{w}_p = Z m_p \quad (7)$$

where $\tilde{A}_p = Z R_p Z^T + \tilde{C}$ and $\tilde{C} = Q^T C Q - Q^T C W F$.

After solving (7) for $\tilde{w}_p$, and using (3) and (6), w_p may be computed as:

$$\begin{aligned} w_p &= Q\tilde{w}_p + W\alpha_p \\ &= (Q - WF)\tilde{w}_p \end{aligned}$$

□