Matrix Inverse Identities

Statistical Machine Learning

Veronica Burt

Woodbury Identity

- We talked about the Woodbury Identity in class, which states for A and C nonsingular,

$$(A - BC^{-1}E)^{-1}BC^{-1} = A^{-1}B(C - EA^{-1}B)^{-1}$$

- Applying this identity to Ridge Regression, we saw

$$\widehat{\beta}_{ridge} = (\mathbb{X}^T\mathbb{X} + \lambda I)^{-1}\mathbb{X}^T\boldsymbol{Y} = \mathbb{X}^T(\mathbb{X}\mathbb{X}^T + \lambda I)^{-1}\boldsymbol{Y}$$

- This results in the inversion of an $n \times n$ matrix as opposed to a $p \times p$ matrix, which can be much less expensive in terms of computation time in the big data setting.

Discovering Matrix Inverse Formulas

- Once a matrix inverse formula is known, it is easy to check that it is true: we just multiply the two matrices together to verify that the result is the Identity Matrix.

- However, discovering the formulas is a much more difficult task.

- Many matrix inverse formulas were discovered by using partitioned (block) matrices.

- Note the following identity for A nonsingular:

$$\begin{bmatrix} A^{-1} & 0 \\ -VA^{-1} & I \end{bmatrix} \begin{bmatrix} A & U \\ V & D \end{bmatrix} = \begin{bmatrix} I & A^{-1}U \\ 0 & D - VA^{-1}U \end{bmatrix}$$

- Using the above equation, we see:

$$\begin{bmatrix} A & U \\ V & D \end{bmatrix}^{-1} = \begin{bmatrix} I & -A^{-1}U(D - VA^{-1}U)^{-1} \\ 0 & (D - VA^{-1}U)^{-1} \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ -VA^{-1} & I \end{bmatrix} =$$

$$\begin{bmatrix} A^{-1} + A^{-1}U(D - VA^{-1}U)^{-1}VA^{-1} & -A^{-1}U(D - VA^{-1}U)^{-1} \\ -(D - VA^{-1}U)^{-1}VA^{-1} & (D - VA^{-1}U)^{-1} \end{bmatrix}$$

- This leads us to the conclusion that inverting a partitioned matrix leads to inverting the sum of two matrices.

- There are many versions of the previous equation, depending on which matrices we require to be nonsingular. For example, here is another version where A, B, U, and V are all nonsingular:

$$\begin{bmatrix} A & U \\ V & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - UD^{-1}V)^{-1} & (V - DU^{-1}A)^{-1} \\ (U - AV^{-1}D)^{-1} & (D - VA^{-1}U)^{-1} \end{bmatrix}$$

Linear Mixed Models

- Consider the linear mixed model,

  $$\boldsymbol{Y} = X\boldsymbol{\beta} + Z\mathbf{u} + \boldsymbol{\epsilon},$$

  where $\boldsymbol{u}$ has dispersion matrix D, independent of $\boldsymbol{\epsilon}$ which has dispersion matrix R.

- This results in $\boldsymbol{Y}$ having the expected value of $X\boldsymbol{\beta}$ and covariance matrix $(R + ZDZ^T)$.

- To find the least squares estimate of $\boldsymbol{\beta}$, we must invert $(R+ZDZ^T)$, an $n \times n$ matrix that often does not have a nice structure(it's normally large and nondiagonal).

- However, the Henderson equations give us

$$(R+ZDZ^T)^{-1} = R^{-1}-R^{-1}Z(Z^TR^{-1}Z+D^{-1})^{-1}Z^TR^{-1},$$

which requires us to invert R, and $n \times n$ matrix that often has a nice structure, D, a $q \times q$ matrix, and $(Z^TR^{-1}Z + D^{-1})$, which is a $q \times q$ matrix.

- Depending on the size of n and q and the structure of D and R, we can choose the formula for $\boldsymbol{\beta}$ which minimizes computation time.

Other Applications

- Intraclass correlation matrices
- Factor analysis
- Discriminant analysis
- Maximum likelihood estimation of variance components

References

Henderson, H.V. and Searle, S. R.

On Deriving the Inverse of a Sum of Matrices

SIAM Review, Vol 23 No 1. 1981.