Matrix Inversion Identities

Ged Ridgway
Centre for Medical Image Computing
University College London

November, 2006

Summary

Two simple matrix identities are derived, these are then used to get expressions for the inverse of $(A + BCD)$. The expressions are variously known as the 'Matrix Inversion Lemma' or 'Sherman-Morrison-Woodbury Identity'.

The derivation in these slides is taken from Henderson and Searle [1]. An alternative derivation, leading to a generalised expression, can be found in Tylavsky and Sohie [2].

Two special case results are mentioned, as they are useful in relating the Kalman-gain form and Information form of the Kalman Filter.

Identity 1

$$
\begin{aligned}
(I + P)^{-1} &= (I + P)^{-1}(I + P - P) \\
&= I - (I + P)^{-1}P
\end{aligned}
\tag{1}
$$

Identity 2

$$P + PQP = P(I + QP) = (I + PQ)P$$
$$(I + PQ)^{-1}P = P(I + QP)^{-1} \tag{2}$$

Matrix Inversion Lemma - step 1

For invertible A, but general (possibly rectangular) B,C, and D:

$$
\begin{aligned}
(A+BCD)^{-1} &= \left(A\left[I + A^{-1}BCD\right]\right)^{-1} \\
&= \left[I + A^{-1}BCD\right]^{-1} A^{-1} \\
&= \left[I - \left(I + A^{-1}BCD\right)^{-1} A^{-1}BCD\right] A^{-1} \qquad \text{Using (1)} \\
&= A^{-1} - (I + A^{-1}BCD)^{-1}A^{-1}BCDA^{-1}
\end{aligned}
$$

Matrix Inversion Lemma - step 2

Repeatedly using (2) in sequence now produces:

$$
\begin{aligned}
(A+BCD)^{-1} &= A^{-1} - (I + A^{-1}BCD)^{-1}A^{-1}BCDA^{-1} && (3)\\
&= A^{-1} - A^{-1}(I + BCDA^{-1})^{-1}BCDA^{-1} && (4)\\
&= A^{-1} - A^{-1}B(I + CDA^{-1}B)^{-1}CDA^{-1} && (5)\\
&= A^{-1} - A^{-1}BC(I + DA^{-1}BC)^{-1}DA^{-1} && (6)\\
&= A^{-1} - A^{-1}BCD(I + A^{-1}BCD)^{-1}A^{-1} && (7)\\
&= A^{-1} - A^{-1}BCDA^{-1}(I + BCDA^{-1})^{-1} && (8)
\end{aligned}
$$

(note that the order ABCD is maintained, ignoring the other parts of the expressions)

Matrix Inversion Lemma - special case

If C is also invertible, from (5):

$$\begin{aligned}(A+BCD)^{-1} &= A^{-1} - A^{-1}B(I+CDA^{-1}B)^{-1}CDA^{-1} \\ &= A^{-1} - A^{-1}B(C^{-1}+DA^{-1}B)^{-1}DA^{-1} \qquad (9)\end{aligned}$$

which is a commonly used variant (for example applicable to the Kalman Filter covariance, in the ‘correction’ step of the filter).

Another related special case

A very similar use of (2) gives:

$$
\begin{aligned}
(A + BCD)^{-1}BC &= A^{-1}(I + BCDA^{-1})^{-1}BC \\
&= A^{-1}B(I + CDA^{-1}B)^{-1}C \\
&\text{and for invertible C:} \qquad (10) \\
&= A^{-1}B(C^{-1} + DA^{-1}B)^{-1} \qquad (11)
\end{aligned}
$$

which is useful in converting between Kalman-gain and Information forms of the Kalman Filter state-estimate 'correction' step.

Bibliography

H. V. Henderson and S. R. Searle.
On Deriving the Inverse of a Sum of Matrices.
SIAM Review, 23(1):53–60, January 1981.

D.J. Tylavsky and G.R.L. Sohie.
Generalization of the matrix inversion lemma.
Proceedings of the IEEE, 74(7):1050–1052, July 1986.