Chaitanya's Random Pages

December 21, 2010

The Matrix Inversion Lemma

Filed under: mathematics — ckrao @ 4:32 am

I had thought the matrix inversion lemma was difficult to prove, but it is in fact not so tricky!

The lemma states that if A and C are square invertible matrices (and B, D are matrices so that A and BCD have the same dimensions), then

$$(A+BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1}+DA^{-1}B)^{-1}DA^{-1} \quad (*)$$

Thanks to [1], it is now easier for me to derive this formula than to remember it, the way much of mathematics should be. Other ways of arriving at the formula are by matrix blockwise elimination or inversion. See the Wikipedia entry on the Woodbury matrix identity (another name for the lemma) for more information.

Proof

1. Start with the equation $(A+BC)x = b$. We find x in terms of b either as $x = (A+BC)^{-1}b$ or as follows.

2. Let $y = Cx$, giving us the two equations

$$\begin{aligned} Ax + By &= b \qquad (1) \\ y &= Cx \qquad (2) \end{aligned}$$

3. From (1) we obtain

$$x = A^{-1}(b - By). \quad (3)$$

4. Substituting (3) into (2) gives $y = CA^{-1}(b - By)$ and rearranging this gives

$$y = (I + CA^{-1}B)^{-1}CA^{-1}b. \quad (4)$$

5. From (3) and (4) we end up with

$$x = A^{-1}b - A^{-1}By = \left[A^{-1} - A^{-1}B(I+CA^{-1}B)^{-1}CA^{-1}\right]b.$$

6. Since b was arbitrary, from step 1 we conclude that

$$(A+BC)^{-1} = A^{-1} - A^{-1}B(I+CA^{-1}B)^{-1}CA^{-1}.$$

7. To arrive at the slightly more complicated form (*) we replace C with CD and note that

$$(I+CDA^{-1}B)^{-1}CD = (I+C\ldots DA^{-1}B)^{-1}D$$

(using the result $(XY)^{-1} = Y^{-1}X^{-1}$).

The matrix inversion lemma is especially use… …, e.g. if they are diagonal or have small dimension. The latter … …h as recursive least squares or the Kalman filter. The lemma is ac… …l inverse theorem, which applies when C is not invertible:

$$(A+BCD)^{-1} = A^{-1} - A\ldots{}^{-1} \quad (*)$$

A couple more special cases of the matrix inv…

1. Sherman-Morrison formula (B and D replaced by column vectors u and v, C replaced by the identity):

$$\left(A+uv^T\right)^{-1} = A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}$$

2.

$$(A+B)^{-1} = A^{-1} - A^{-1}B\left(B+BA^{-1}B\right)^{-1}BA^{-1}$$

3.

$$(I+A)^{-1} = I - (I+A)^{-1}A$$

Reference:

[1] S. Boyd and L. Vandenberghe, Convex Optimization (Appendix C.4.3), Cambridge University Press, 2004

Share this:

Like

One blogger likes this.

Related

Inverse variance weighting form of the conditional covariance of multivariate Gaussian vectors
In "mathematics"

An easier way to square some 2x2 matrices
In "mathematics"

Some notes on the Schur Complement
In "mathematics"

Comments (1)

1 Comment »

1. nice proof

   Comment by ekaveera — May 21, 2012 @ 2:54 pm | Reply

RSS feed for comments on this post. TrackBack URI

Leave a Reply

Enter your comment here...

- Pages
  - About
- Categories:
  - climate and weather
  - cricket
  - geography
  - mathematics
  - movies and TV
  - nature
  - science
  - sport

- Uncategorized
- Search:

  Search
- Archives:
  - August 2014
  - July 2014
  - June 2014
  - May 2014
  - April 2014
  - March 2014
  - February 2014
  - January 2014
  - December 2013
  - November 2013
  - October 2013
  - September 2013
  - August 2013
  - July 2013
  - June 2013
  - May 2013
  - April 2013
  - March 2013
  - February 2013
  - January 2013
  - December 2012
  - November 2012
  - October 2012
  - September 2012
  - August 2012
  - July 2012
  - June 2012
  - May 2012
  - April 2012
  - March 2012
  - February 2012
  - January 2012
  - December 2011
  - November 2011
  - October 2011
  - September 2011
  - August 2011
  - July 2011
  - June 2011
  - May 2011
  - April 2011
  - March 2011
  - February 2011
  - January 2011
  - December 2010

- Meta:
  - Register
  - Log in
  - RSS
  - Comments RSS
  - XFN
  - Create a free website or blog at WordPress.com.