

Inverses of tridiagonal matrices

Gérard MEURANT

<https://gerard-meurant.fr>

Selva di Fasano, August 29-September 1, 2023

Let us start with unreduced Hessenberg matrices H of order n ($h_{j+1,j} \neq 0, j = 1, \dots, n-1$)

$$H = UCU^{-1}$$

where U is a nonsingular upper triangular matrix of order n with $u_{1,1} = 1$ and diagonal entries $u_{k,k} = \prod_{j=1}^{k-1} h_{j+1,j}$ for $k > 1$ and

$$C = \begin{pmatrix} 0 & 0 & \dots & 0 & -\alpha_0 \\ 1 & 0 & \dots & 0 & -\alpha_1 \\ 0 & 1 & \dots & 0 & -\alpha_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 1 & -\alpha_{n-1} \end{pmatrix}$$

where

$$p(\lambda) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \dots + \alpha_1\lambda + \alpha_0$$

is the characteristic polynomial of H

To prove this, set $U = (e_1 \quad He_1 \quad H^2e_1 \quad \dots \quad H^{n-1}e_1)$ and use the Cayley-Hamilton theorem which gives $HU = UC$

We are interested in H^{-1} . Let

$$\beta_1 = -\alpha_1/\alpha_0, \quad \hat{\beta} = -(\alpha_2/\alpha_0 \quad \cdots \quad \alpha_{n-1}/\alpha_0 \quad 1/\alpha_0)^T,$$

$$U^{-1} = \begin{pmatrix} 1 & \hat{\vartheta}^T \\ 0 & \hat{U}^{-1} \end{pmatrix}.$$

Then, the inverse of H can be written as

$$H^{-1} = \begin{pmatrix} \beta_1 - \hat{\vartheta}^T \hat{U} \hat{\beta} & (\beta_1 - \hat{\vartheta}^T \hat{U} \hat{\beta}) \hat{\vartheta}^T + (e_1^T - \hat{\vartheta}^T \hat{U} F) \hat{U}^{-1} \\ \hat{U} \hat{\beta} & \hat{U} \hat{\beta} \hat{\vartheta}^T + \hat{U} F \hat{U}^{-1} \end{pmatrix},$$

where F is the zero matrix except for the entries on the first upper diagonal which are equal to 1

Proof: $H^{-1} = UC^{-1}U^{-1}$

$$C^{-1} = \begin{pmatrix} \beta_1 & \mathbf{e}_1^T \\ \hat{\beta} & F \end{pmatrix}$$

The inverse of the matrix U^{-1} is

$$U = \begin{pmatrix} 1 & -\hat{\vartheta}^T \hat{U} \\ 0 & \hat{U} \end{pmatrix}$$

The matrix $\hat{U}F\hat{U}^{-1}$ is strictly upper triangular

Therefore, the lower triangular part of the principal trailing block of H^{-1} is the lower triangular part of $\hat{U}\hat{\beta}\hat{\vartheta}^T$. We easily see that the lower triangular part of H^{-1} is the lower triangular part of a rank-one matrix

This was proved by Y. Ikebe (1979) and D.K. Fadjev (1984) with different proofs

Let us apply these results to an unreduced symmetric tridiagonal matrix T

$$T = \begin{pmatrix} \alpha_1 & \beta_2 & & & \\ \beta_2 & \alpha_2 & \beta_3 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{n-1} & \alpha_{n-1} & \beta_n \\ & & & \beta_n & \alpha_n \end{pmatrix}$$

with $\beta_j \neq 0, j = 2, \dots, n$

A symmetric upper Hessenberg matrix is symmetric tridiagonal

Hence, the lower triangular part of T^{-1} is the lower triangular part of a rank-one matrix

Because of the symmetry, T^{-1} has the following structure

There exist two sequences of nonzero numbers $\{\nu_i\}, \{\sigma_i\}, i = 1, \dots, n$ with $\nu_1 = 1$ such that

$$T^{-1} = \begin{pmatrix} \nu_1\sigma_1 & \nu_1\sigma_2 & \nu_1\sigma_3 & \dots & \nu_1\sigma_n \\ \nu_1\sigma_2 & \nu_2\sigma_2 & \nu_2\sigma_3 & \dots & \nu_2\sigma_n \\ \nu_1\sigma_3 & \nu_2\sigma_3 & \nu_3\sigma_3 & \dots & \nu_3\sigma_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \nu_1\sigma_n & \nu_2\sigma_n & \nu_3\sigma_n & \dots & \nu_n\sigma_n \end{pmatrix}$$

This type of matrices were called *matrices factorisables* by J. Baranger and M. Duc-Jacquet (1971) and *symmetric semiseparable* or *symmetric generator representable semiseparable matrices* by R. Vandebril, M. Van Barel, and N. Mastronardi (2008)¹

¹Matrix Computations and Semiseparable Matrices, Volume I, The Johns Hopkins University Press

T^{-1} being nonsingular if and only if $\nu_i \neq 0, i = 1, \dots, n$ and $\nu_i \sigma_{i+1} - \nu_{i+1} \sigma_i \neq 0, i = 1, \dots, n - 1$

The relation

$$\frac{\sigma_i}{\sigma_{i+1}} > \frac{\nu_i}{\nu_{i+1}}, \quad i = 1, \dots, n - 1$$

corresponds to T^{-1} (and T) being positive definite

We are interested in the inverses of the principal submatrices $T_k, k = 1, \dots, n - 1$ of T

Let us start with T_{n-1} . We have to remove the last row and the last column of T

$$T_{n-1}^{-1} = \begin{pmatrix} \nu_1^{(n-1)} \sigma_1^{(n-1)} & \nu_1^{(n-1)} \sigma_2^{(n-1)} & \nu_1^{(n-1)} \sigma_3^{(n-1)} & \dots & \nu_1^{(n-1)} \sigma_n^{(n-1)} \\ \nu_1^{(n-1)} \sigma_2^{(n-1)} & \nu_2^{(n-1)} \sigma_2^{(n-1)} & \nu_2^{(n-1)} \sigma_3^{(n-1)} & \dots & \nu_2^{(n-1)} \sigma_n^{(n-1)} \\ \nu_1^{(n-1)} \sigma_3^{(n-1)} & \nu_2^{(n-1)} \sigma_3^{(n-1)} & \nu_3^{(n-1)} \sigma_3^{(n-1)} & \dots & \nu_3^{(n-1)} \sigma_n^{(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \nu_1^{(n-1)} \sigma_n^{(n-1)} & \nu_2^{(n-1)} \sigma_n^{(n-1)} & \nu_3^{(n-1)} \sigma_n^{(n-1)} & \dots & \nu_n^{(n-1)} \sigma_n^{(n-1)} \end{pmatrix}$$

Let \mathcal{B} be a nonsingular matrix of order n , partitioned as

$$\mathcal{B} = \begin{pmatrix} B & a \\ c^T & \alpha \end{pmatrix}$$

with B nonsingular of order $n - 1$

We define two vectors

$$u = \mathcal{B}_{:,n} - e_n = \begin{pmatrix} a \\ \alpha - 1 \end{pmatrix}, \quad v = e_n$$

and the rank-one modification

$$\mathcal{B} - uv^T = \begin{pmatrix} B & 0 \\ c^T & 1 \end{pmatrix}$$

Using the Sherman-Morrison formula

$$(\mathcal{B} - uv^T)^{-1} = \mathcal{B}^{-1} + \frac{1}{[\mathcal{B}^{-1}]_{n,n}} [e_n - [\mathcal{B}^{-1}]_{:,n}] [[\mathcal{B}^{-1}]_{n,:}].$$

If we know the inverse of \mathcal{B} , by taking the indices (i, j) such that $1 \leq i, j \leq n - 1$ in the right-hand side, we obtain the inverse of B

To compute T_{n-1}^{-1} , we apply the previous results with $\mathcal{B} = T$

We just have to look at the first and last columns of the inverse

For the first column

$$\begin{pmatrix} \sigma_1^{(n-1)} \\ \sigma_2^{(n-1)} \\ \vdots \\ \sigma_{n-1}^{(n-1)} \end{pmatrix} = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_{n-1} \end{pmatrix} - \frac{1}{\nu_n} \begin{pmatrix} \sigma_n \\ \sigma_n \nu_2 \\ \vdots \\ \sigma_n \nu_{n-1} \end{pmatrix}$$

It yields

$$\sigma_i^{(n-1)} = \sigma_i - \sigma_n \frac{\nu_i}{\nu_n}, \quad i = 1, 2, \dots, n-1$$

For the last column

$$\sigma_{n-1}^{(n-1)} \begin{pmatrix} 1 \\ \nu_2^{(n-1)} \\ \vdots \\ \nu_{n-1}^{(n-1)} \end{pmatrix} = \sigma_{n-1} \begin{pmatrix} 1 \\ \nu_2 \\ \vdots \\ \nu_{n-1} \end{pmatrix} - \frac{\nu_{n-1}}{\nu_n} \begin{pmatrix} \sigma_n \\ \sigma_n \nu_2 \\ \vdots \\ \sigma_n \nu_{n-1} \end{pmatrix}$$

It yields

$$\sigma_{n-1}^{(n-1)} \nu_i^{(n-1)} = \sigma_{n-1} \nu_i - \sigma_n \frac{\nu_{n-1}}{\nu_n} \nu_i, \quad 1 \leq i \leq n-1.$$

Using the value of $\sigma_{n-1}^{(n-1)}$,

$$\sigma_{n-1}^{(n-1)} \nu_i^{(n-1)} = \sigma_{n-1} \nu_i - \sigma_n \frac{\nu_{n-1}}{\nu_n} \nu_i = \left(\sigma_{n-1} - \sigma_n \frac{\nu_{n-1}}{\nu_n} \right) \nu_i^{(n-1)}.$$

Since $\sigma_{n-1}^{(n-1)} \neq 0$, it yields $\nu_i^{(n-1)} = \nu_i$, for $1 \leq i \leq n-1$

To obtain the inverse of T_k , with $k = 1, \dots, n - 2$ we apply the previous result recursively

$$T_k^{-1} = \begin{pmatrix} \sigma_1^{(k)} & \sigma_2^{(k)} & \cdots & \sigma_k^{(k)} \\ \sigma_2^{(k)} & \sigma_2^{(k)} \nu_2 & \cdots & \sigma_k^{(k)} \nu_2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_k^{(k)} & \sigma_k^{(k)} \nu_2 & \cdots & \sigma_k^{(k)} \nu_k \end{pmatrix}$$

with

$$\sigma_i^{(k)} = \sigma_i^{(k+1)} - \sigma_{k+1}^{(k+1)} \frac{\nu_i}{\nu_{k+1}}, \quad i = 1, 2, \dots, k$$

Moreover

$$\sigma_i^{(k)} = \sigma_i - \sigma_{k+1} \frac{\nu_i}{\nu_{k+1}}, \quad i = 1, 2, \dots, k$$

The last relation is proved by induction

We can apply these results to the symmetric tridiagonal matrices generated by the Lanczos algorithm and implicitly by CG

Let $\nu^T = (1 \ \nu_2 \ \cdots \ \nu_n)$. The last column of T^{-1} is proportional to ν

$$T^{-1}e_n = \sigma_n \nu \Rightarrow \sigma_n \nu^T T = e_n^T$$

We have $T = UCU^{-1}$ and

$$\nu^T U = (1 \ \nu^T T e_1 \ \nu^T T^2 e_1 \ \cdots \ \nu^T T^{n-1} e_1) = e_1^T$$

ν^T is equal to the first row of U^{-1}

Let x_k and r_k be the CG iterates and residual vectors when solving $Ax = b$. From the results of G.M. and J. Duintjer Tebbens², we can show that

$$\nu_{k+1} = (-1)^k \frac{\|r_0\|}{\|r_k\|}, \quad \sigma_{k+1} = (-1)^k \frac{\|x - x_k\|_A^2}{\|r_0\| \|r_k\|}$$

$$\sigma_i^{(k)} = (-1)^{i-1} \frac{1}{\|r_0\| \|r_{i-1}\|} \sum_{j=i-1}^k \gamma_j \|r_j\|^2, \quad i = 1, \dots, k$$

where

$$\gamma_j = \frac{r_j^T r_j}{p_j^T A p_j}$$

is one of the CG coefficients

We conclude with a quote from Cornelius Lanczos

Retirement merely means one carries on, on half pay

Nevertheless, HAPPY RETIREMENT Marc