On Pascal-like matrices for the compression and reconstruction of sparse Jacobian

Shahadat Hossain
Department of Mathematics and Computer Science
University of Lethbridge, Canada
and
Institut für Mathematik, Humboldt-Universität
10099 Berlin, Germany

3rd European Workshop on Automatic Differentiation
Oxford University Computing Laboratory, Oxford, UK
June 01, 2006

Outline

- The Jacobian matrix determination problem
- Compression and reconstruction
- Elementary merges and Pascal seeding
- Linear algebra
- Pascal and Vandermonde connection
- Concluding remarks

The Problem

$$\left.\frac{\partial F(x+ts)}{\partial t}\right|_{t=0} = F'(x)s \approx As = \frac{1}{\varepsilon}[F(x+\varepsilon s) - F(x)] \equiv b$$

with a forward difference (one extra function evaluation) gives $As = b$ where b is the finite difference approximation.

Algorithmic Differentiation (AD) gives $b = F'(x)s$.

Formulation of the problem: Obtain vectors $s_1, \cdots, s_p$ such that the matrix vector product

$$b_i \equiv As_i,\ i = 1, \cdots, p \text{ or } B \equiv AS$$

determine the $m \times n$ matrix A uniquely.

Main Steps in Computing A (Notations)

A: $m \times n$ matrix to be determined

ρ_i: Number of nonzero entries in row i of A

v_i: vector of column indices of nonzero entries in row i of A

S: $n \times p$ “seed” matrix is given

Obtain $B = AS$ (p matrix-vector products).

Main Steps in Computing A (Procedure)

Assume that $\rho_i \leq p$ for all i

Determine the nonzero elements of A row by row:

1. Identify the reduced seed matrix $\hat{S}_i \in \mathbf{R}^{\rho_i \times p}$ for $A(i, v_i)$

   (the ρ_i nonzeros in row i of A)

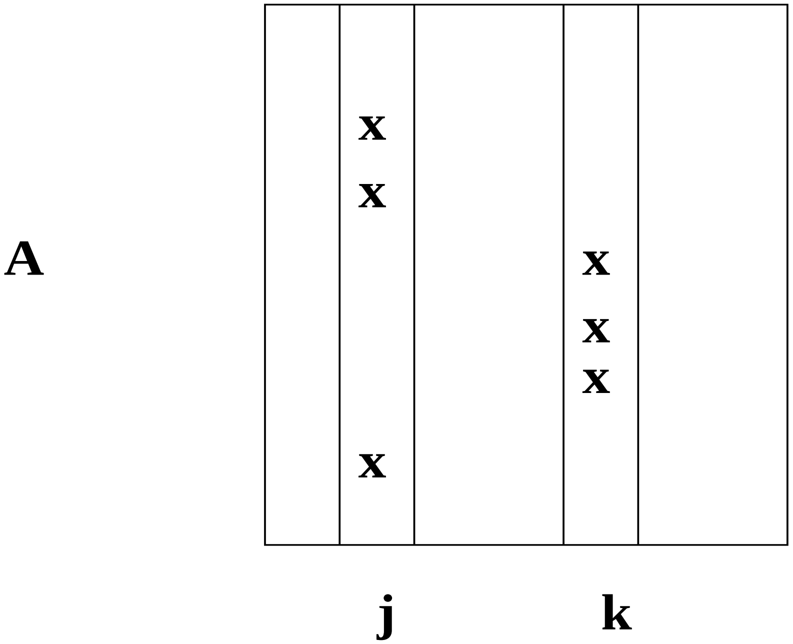
$$\hat{S}_i = S(v_i, :)$$

2. Solve for unknown elements $a_{ik} \neq 0$ of $A(i, :)$

$$\hat{S}_i{}^T A(i, v_i)^T = B(i, :)^T$$

Curtis, Powell, and Reid (1974)

$$F'_j + F'_k \approx a_j + a_k = \tfrac{1}{\varepsilon}[F(x+\varepsilon(e_j+e_k)) - F(x)]$$



Determine the elements in columns j and k in the $m \times n$ matrix A from $As = b$ where

$$s = e_j + e_k, \text{ and } b = \frac{1}{\varepsilon}[F(x+\varepsilon(e_j+e_k)) - F(x)]$$

Direct Determination

$$A = \begin{bmatrix} a_{11} & 0 & a_{13} \\ a_{21} & a_{22} & 0 \\ 0 & a_{32} & a_{33} \end{bmatrix}$$

Obtaining the nonzeros explicitly requires three matrix vector products.

- Choose $S = I$ and obtain

  $$B = AS$$

  using three matrix vector products. The nonzeros are read off directly (without any other arithmetic operations) from B.

Indirect Determination

$$A = \begin{bmatrix} a_{11} & 0 & a_{13} \\ a_{21} & a_{22} & 0 \\ 0 & a_{32} & a_{33} \end{bmatrix}$$

The nonzeros can be determined implicitly by a substitution process.

- Choose

$$S = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

and obtain

$$B = AS$$

requiring only “two” matrix-vector products and “three” 2×2 triangular linear system solves.

Determine row 2 of A

$$A = \begin{bmatrix} a_{11} & 0 & a_{13} \\ a_{21} & a_{22} & 0 \\ 0 & a_{32} & a_{33} \end{bmatrix}$$

$$\begin{bmatrix} a_{21} & a_{22} & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} b_{21} & b_{22} \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{21} \\ a_{22} \end{bmatrix} = \begin{bmatrix} b_{21} \\ b_{22} \end{bmatrix}$$

Seeding for compression

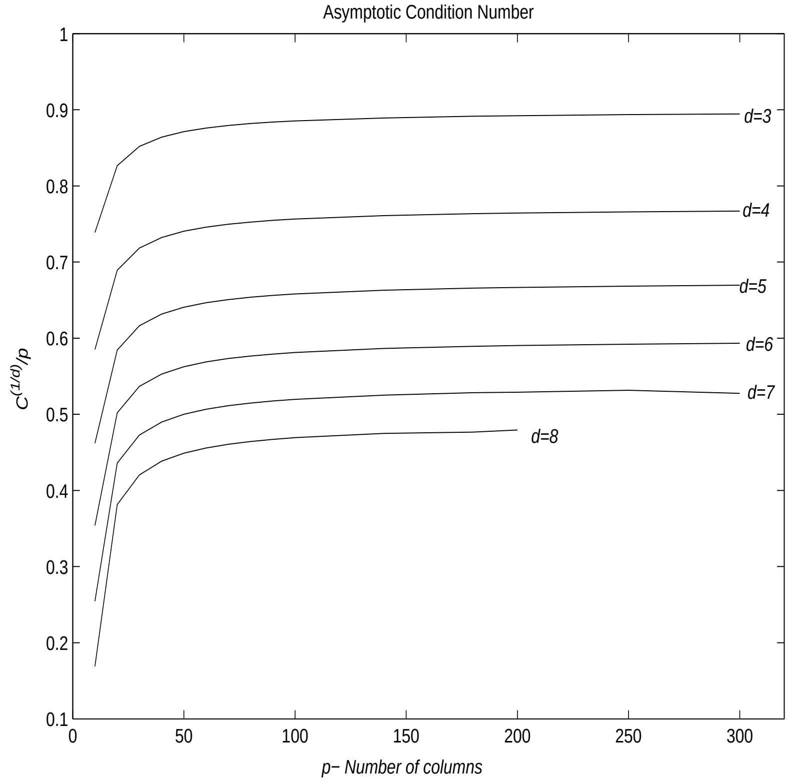
Find suitable seed matrices any square submatrix of which is *numerically well-conditioned* and *easy to solve*

1. Direct Methods
   - (a) The Curtis, Powell, and Reid (CPR) [1974]
   - (b) Powell and Toint [1979], Coleman and Moré [1983]
   - (c) Goldfarb and Toint [1984]
   - (d) Hossain and Steihaug [1998], Coleman and Verma [1998]
   - (e) Hossain and Steihaug [2003], Gebremedhin, Manne, Pothen [2004]

1. Indirect Methods
   (a) Vandermonde Seeding: (Newsam and Ramsdell [1983], Geitner, Utke and, Griewank [1996])
   (b) Column merging seed matrices proposed in Hossain and Steihaug [AD2000] yields determination by substitution
   (c) Pascal Seeding: (Hossain and Steihaug [2002, 2004]), Newsam-Ramsdell-Lagrange (Griewank and Verma [2002])
   Let 0_j denote the zero vector in $\mathbf{R}^j$. Then column i in the seed matrix S is (row i in S^T)
   $$\left[0_{i-1}\; u\; 0_{p-d-i}\right]$$
   where component j in $u \in \mathbf{R}^{d+1}$ is the binomial coefficient $\binom{d}{j-1}$
   (The growth of condition number is $O(p^d)$.)

‘**Numerical Test Results** 1

Figure 1: Condition estimate for Pascal Seed matrices



Pascal Seeding

$$\underset{(a)}{S^{(3)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 3 & 3 & 1 & 0 & 0 \\ 1 & 3 & 3 & 1 & 0 \\ 0 & 1 & 3 & 3 & 1 \\ 0 & 0 & 1 & 3 & 3 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}}, \quad \underset{(b)}{a = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}}, \quad \underset{(c)}{a^T S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 3 & 1 & 0 & 0 \\ 0 & 1 & 3 & 3 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}}$$

Figure 2: (a) Pascal matrix for $p = 8$ and $d = 3$; (c) A submatrix $a^T S^{(3)}$ consisting of rows 1, 3, 5, and 8

Pascal matrices - Construction

Define the *elementary merge matrix*

$$\begin{aligned} &E_n(i,i) = E_n(i+1,i) = 1 && \text{for } i = 1,2,\ldots,n-1 \\ &E_n(i,j) = 0 && \text{if } j > i \text{ or } j < i-1 \end{aligned}$$

Results:

1. The Pascal matrix $S^{(n-\rho)} \in \mathbb{R}^{n\times\rho}, n > \rho \geq 1$ can be factorized into $n-\rho$ elementary merge matrices

$$S^{(n-\rho)} = E_n E_{n-1} \ldots E_{\rho+1}$$

2. For $n > \rho+1 > 2$

$$S^{(n-\rho)} = S^{(n-\rho-1)} E_{\rho+1}.$$

Example

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 3 & 3 & 1 \\ 1 & 3 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}}_{S^{(3)}} = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}}_{E_5} * \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{E_4} * \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}}_{E_3}$$

Calculating the Inverse-I

Let

$$\hat{S}^{(\rho)} \equiv S^{(n-\rho)}\left([1\ldots\rho],:\right),$$

denote the triangular submatrix consisting of first ρ rows of S. Set $d \equiv n-\rho$.

Theorem. Given the Pascal-r matrix $S^{(n-\rho)} \in \mathbb{R}^{n\times\rho}$, we have

$$\hat{S}^{(\rho)}{}^{-1}_{ij} = \begin{cases} (-1)^{i+j}\binom{d+i-j-1}{d-1} & \text{for } i \geq j \\ 0 & \text{for } i < j \end{cases} \tag{1}$$

$\square$

Calculating the Inverse-II

Lemma 1 *Every square submatrix* $\hat{E_n}$ *of* $E_n, n \geq 3$ *which is not the leading (first* $n-1$ *rows) or the trailing (last* $n-1$ *rows) is block diagonal and has the form*

$$\hat{E_n} = \begin{pmatrix} \hat{E_k} & 0 \\ 0 & \hat{E_l} \end{pmatrix}$$

where $k+l = n-2$ *with* $2 \leq k, l \leq n-1$ *and* $\hat{E_k}$ *and* $\hat{E_l}$ *are leading and trailing square submatrices of* E_k *and* E_l *respectively.*

Theorem. Every square submatrix of E_n for $n \geq 2$ is invertible. ☐

Lemma 2 *Let* $T^{(n)} \in \mathbb{R}^{n\times n}$ *be a tridiagonal matrix defined as*

$$a_{i,j} = \begin{cases} 1 & \text{for } i+1=j, \text{ or } i-1=j, \\ 2 & \text{for } i=j, \text{ and,} \\ 0 & \text{for } |i-j|>1. \end{cases} ,$$

then

$$A^{-1} = diag\left(\frac{1}{n+1},\ldots,\frac{1}{n+1}\right)B$$

where B is symmetric and is given by

$$b_{ij} = \begin{cases} (-1)^{i+j}(n-i+1)j & \text{for } j \le i \\ (-1)^{i+j}(n-j+1)i & \text{for } j \ge i+1 \end{cases}$$

$i=1,\ldots,n,\ j=1,\ldots,n.$

The Vandermonde Connection:Recurrence Relation

Let $B = A^{-1}$ be the inverse of the band Toeplitz matrix

$$A = \begin{pmatrix} a_0 & a_1 & \cdots & a_u & & & & & \\ a_{-1} & a_0 & a_1 & \cdots & a_u & & & & \\ \vdots & \ddots & \ddots & & & & & & \\ a_{-l} & a_{-l+1} & \cdots & a_0 & \cdots & & a_u & & \\ & & & & \ddots & & & \ddots & \\ & & a_{-l} & a_{-l+1} & \cdots & a_0 & \cdots & a_{u-1} & a_u \\ & & & & & & \ddots & & \vdots \\ & & & & a_{-l} & a_{-l+1} & \cdots & a_0 & a_1 \\ & & & & & a_{-l} & \cdots & a_{-1} & a_0 \end{pmatrix},$$

of dimension $n \times n$ where $l, u \geq 0$ are the lower and the upper bandwidth, respectively and the nonzero entries are the binomial coefficients $a_k = \binom{l+u}{l+k}, k = -l \ldots 0 \ldots u$. Using $AB = I$, the entries b_{ij} satisfies

$$\sum_{k=-l}^{u} a_k b_{k+i,j} = 0 \quad \text{if } i \neq j \tag{2}$$

$$\sum_{k=-l}^{u} a_k b_{k+j,j} = 1 \quad \text{if } i = j. \tag{3}$$

The General Solution

$$b_{ij} = \begin{cases} (-1)^i \sum_{k=0}^{l+u} c_k(j) i^k, & 1 \le i \le j+u-1 \\ (-1)^i \sum_{k=0}^{l+u} [c_k(j) - d_k(j)] i^k, & j-l+1 \le i \le n \end{cases} \tag{4}$$

Determine the constants c_k and d_k from the boundary conditions and equation (3).

Solving for the constants

$$
\begin{aligned}
b_{ij} &= 0 \qquad \text{(5a)}\\
\text{for } i &= 1-l,\ldots,0 \text{ if } l \geq 1\\
\text{for } i &= n,\ldots,n+u \text{ if } u \geq 1\\
(-1)^i \sum_{k=0}^{l+u} c_k(j) i^k &= (-1)^i \sum_{k=0}^{l+u} [c_k(j) - d_k(j)] i^k \qquad \text{(5b)}\\
\text{for } i &= j-l+1 \ldots j+u-1 \text{ if } l+u > 1\\
-(-1)^{j+u} \sum_{k=0}^{l+u} c_k(j)(j+u)^k + (-1)^{j+q} \sum_{k=0}^{l+u} [c_k(j) - d_k(j)](j+u)^k &= 1/a_u \qquad \text{(5c)}\\
\text{if } l+u &> 1.
\end{aligned}
$$

Solving for the constants

Solve for d_k:

$$\sum_{k=1}^{l+u}(-1)^{j-l+q}(j-l+q)^k d_k(j) = \begin{cases} 0 & q = 1\ldots l+u-1 \\ -1/a_u & q = l+u \end{cases} \tag{6}$$

Solve for c_k:

$$\sum_{k=1}^{l+u}(-1)^q q^k c_k(j) = \begin{cases} 0 & q = 1-l\ldots 0 \\ \sum_{k=1}^{l+u}(-1)^q q^k c_k(j) & q = n+1\ldots n+u \end{cases} \tag{7}$$

Concluding Remarks

- The leading/trailing triangles are the worst-conditioned reduced system
- The Pascal seeding elimination provides a flexible solution
  - For practical purposes d is small and the condition number of the seed matrix is of the order p^d (polynomial for practical purposes)
  - The linear system solves can be implemented in qd^2 (linear in q for practical purposes)
  - Accuracy and time complexity can be adapted to the user requirement
  - The generation, storage, and manipulation of the seed matrix is efficient and convenient