Gram-Schmidt Orthogonalization: 100 Years and More

Steven Leon, Åke Björck, Walter Gander

April 4, 2009
Outline of Paper

- Early History
- CGS and MGS and QR
- Elimination Methods and MGS
- Reorthogonalization
- Rank Revealing factorizations
- Gram-Schmidt in Iterative Methods
- Implementing Gram-Schmidt Algorithms
Outline of Talk

- Early History (1795–1907)

- Middle History
  1. The work of Åke Björck
     Least squares, Stability, Loss of orthogonality
  2. The work of Heinz Rutishauser
     Selective reorthogonalization and
     Superorthogonalization

Acknowledgement: Thanks to Julien Langou for providing his analysis of Laplace’s work on MGS and for proof reading and suggesting improvements in my slides.
Early History

- Least Squares - Gauss and Legendre
- Laplace 1812, Analytic Theory of Probabilities (1814, 1820)
- Cauchy 1837, 1847 (Interpolation) and Bienaymé 1853
- J. P. Gram, 1879 (Danish), 1883 (German)
- Erhard Schmidt 1907
Gauss and least squares

- Priority dispute:
  A. M. Legendre - first publication 1805;
  Gauss claimed discovery in 1795

- G. Piazzi, January 1801 discovered asteroid Ceres.
  Tracked it for 6 weeks.
The Gauss derivation of the normal equations

Let $\hat{x}$ be the solution $A^T Ax = A^T b$.
For any $x$ we have

$$r = (b - A\hat{x}) + A(\hat{x} - x) \equiv \hat{r} + Ae,$$

and since $A^T \hat{r} = A^T b - A^T A\hat{x} = 0$

$$r^T r = (\hat{r} + Ae)^T (\hat{r} + Ae) = \hat{r}^T \hat{r} + (Ae)^T (Ae).$$

Hence $r^T r$ is minimized when $x = \hat{x}$. 
Pierre-Simon Laplace (1749–1827)

- Mathematical Astronomy
- Celestial Mechanics
- Laplace’s Equation
- Laplace Transforms
- Probability Theory and Least Squares
Problem: Compute masses of Saturn and Jupiter from systems of normal equations (Bouvart) and to compute the distribution of error in the solutions.

Method: Laplace successively projects the system of equations orthogonally to a column of the observation matrix to eliminate all variables but the one of interest.

Basically Laplace uses MGS to prove that, given an overdetermined system with a normal perturbation on the right-hand side, its solution has a centered normal distribution with variance independent from the parameters of the noise.
Laplace uses MGS to derive the Cholesky form of the normal equations, $R^T R \mathbf{x} = A^T \mathbf{x}$

Laplace does not seem to realize that the vectors generated are mutually orthogonal.

He does observe that the generated vectors are each orthogonal to the residual vector.
Cauchy (1837) and (1847)- interpolation method leading to systems of the form $Z^T A x = Z^T b$, where $Z = (z_1, \ldots, z_n)$ and $z_{ij} = \pm 1$.

Bienaymé (1853) new derivation of Cauchy’s algorithm based on Gaussian elimination.

Bienaymé noted that the Cauchy’s choice of $Z$ was not optimal in the least squares sense. Least squares solution if $Z = A$ (normal equations) or more generally if $R(Z) = R(A)$. The matrix $Z$ can be determined a column at a time as the elimination steps are carried out.
\[ \begin{pmatrix} z_1^T a_1 & z_2^T a_2 & z_3^T a_3 \\ z_2^T a_1 & z_2^T a_2 & z_2^T a_3 \\ z_3^T a_1 & z_3^T a_2 & z_3^T a_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} z_1^T b \\ z_2^T b \\ z_3^T b \end{pmatrix} \]

Transform \( i,j \)th element (\( 2 \leq i, j \leq 3 \))

\[ z_i^T a_j - \frac{z_i^T a_1}{z_1^T a_1} z_1^T a_j = z_i^T \left( a_j - \frac{z_1^T a_j}{z_1^T a_1} a_1 \right) \equiv z_i^T a_j^{(2)}, \]
The reduced system has the form

\[
\begin{pmatrix}
  z_2^T a_2^{(2)} & z_2^T a_3^{(2)} \\
z_3^T a_2^{(2)} & z_3^T a_3^{(2)}
\end{pmatrix}
\begin{pmatrix}
  x_2 \\
x_3
\end{pmatrix}
= \begin{pmatrix}
  z_2^T b^{(2)} \\
z_3^T b^{(2)}
\end{pmatrix}
\]

where we have defined

\[
a_j^{(2)} = a_j - \frac{z_1^T a_j}{z_1^T a_1} a_1, \quad b^{(2)} = b - \frac{z_1^T b}{z_1^T b} a_1.
\]

Finally \(z_3\) is chosen and used to form the single equation

\[
z_3^T a_3^{(3)} x_3 = z_3^T b^{(3)}.
\]

Taking the first equation from each step gives a triangular system defining the solution.
An interesting choice of $Z$

Since we want $R(Z) = R(A)$, if we choose $Z = Q$ where

$$q_1 = a_1, \quad q_2 = a_2^{(2)}, \quad q_3 = a_3^{(3)}, \ldots$$

Then we have

$$q_2 = a_2 - \frac{q_1^T a_2}{q_1^T q_1} q_1, \quad q_3 = a_3^{(2)} - \frac{q_2^T a_3^{(2)}}{q_2^T q_2} q_2, \ldots,$$

which is exactly the **modified Gram-Schmidt** procedure!
Jørgen Pedersen Gram (1850–1916)

- Dual career: Mathematics (1873) and Insurance (1875, 1884)
- Research in modern algebra, number theory, models for forest management, integral equations, probability, numerical analysis.
- Active in Danish Math Society, edited Tidsskrift journal (1883–89).
- Best known for his orthogonalization process.
J. P. Gram: 1879 Thesis and 1883 paper

- Series expansions of real functions using least squares

- Orthogonalization process applied to generate orthogonal polynomials

- Data approximation using discrete inner product

- Determinantal representation for resulting orthogonal functions

- Continuous inner products

- Application to integral equations
Erhard Schmidt (1876–1959)

- Ph.D. on integral equations, Gottengin 1905
- Student of David Hilbert
- 1917 University of Berlin, set up Inst for Applied Math
- Director of Math Res Inst of German Acad of Sci
- Known for his work in Hilbert Spaces
- Played important role in the development of modern functional analysis
p 442 CGS for sequence of functions $\phi_1, \ldots, \phi_n$ with respect to inner product

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx$$

p 473 CGS for an infinite sequence of functions

In footnote (p 442) Schmidt claims that in essence the formulas are due to J. P. Gram.

1935 Y. K. Wong paper refers to “Gram-Schmidt Orthogonalization Process” (First such linkage?)
Original algorithm of Schmidt for CGS

\begin{align*}
\psi_1(x) &= \frac{\phi_1(x)}{\sqrt{\int_a^b \phi_1(y)^2 dy}} \\
\psi_2(x) &= \frac{\phi_2(x) - \psi_1(x) \int_a^b \phi_2(z) \psi_1(z) dz}{\sqrt{\int_a^b (\phi_2(y) - \psi_1(y) \int_a^b \phi_2(z) \psi_1(z) dz)^2 dy}} \\
&\quad \vdots \\
\psi_n(x) &= \frac{\phi_n(x) - \sum_{\rho=1}^{n-1} \psi_\rho(x) \int_a^b \phi_n(z) \psi_\rho(z) dz}{\sqrt{\int_a^b (\phi_n(x) - \sum_{\rho=1}^{n-1} \psi_\rho(x) \int_a^b \phi_n(z) \psi_\rho(z) dz)^2 dy}}
\end{align*}
Åke Björck

- Specialist in Numerical Least Squares
- 1967 Solving Least Squares Problems by Gram-Schmidt Orthogonalization
- 1992 Björck and Paige: Loss and recapture of orthogonality in MGS
- 1994 Numerics of Gram-Schmidt Orthogonalization
- 1996 SIAM: Numerical Methods for Least Squares
Stability of MGS for Least Squares

- **Forward Stability** (Björck, 1967) and **Backward Stability** (Björck and Paige, 1992)

- If $A$ has computed MGS factorization $\tilde{Q}\tilde{R}$
  - Loss of orthogonality in MGS

\[
\|I - \tilde{Q}^T\tilde{Q}\|_2 \leq \frac{c_1(m, n)}{1 - c_2(m, n)\kappa u}
\]

- **Stability**

\[
A + E = Q\tilde{R} \text{ where } \|E\| \leq c(m, n)u\|A\|_2 \text{ and } Q^TQ = I
\]

- $b$ must be modified as if it were an $n + 1$st column of $A$
\[ \tilde{A} = \begin{pmatrix} O \\ \tilde{A} \end{pmatrix} = \tilde{Q} \tilde{R} = \begin{pmatrix} \tilde{Q}_{11} & \tilde{Q}_{12} \\ \tilde{Q}_{21} & \tilde{Q}_{22} \end{pmatrix} \begin{pmatrix} \tilde{R}_1 \\ O \end{pmatrix} \]

where \( \tilde{Q} = H_1 H_2 \cdots H_n \) (a product of Householder matrices)

\[ \tilde{Q}_{11} = O \text{ and } A = \tilde{Q}_{21} \tilde{R}_1 \]

(C. Sheffield) \( \tilde{Q}_{21} \tilde{R}_1 \) and \( A = QR \) (MGS) \textit{numerically} equivalent

In fact if \( Q = (q_1, \ldots, q_n) \) then

\[ H_k = I - v_k v_k^T, \quad k = 1, \ldots, n \]

where

\[ v_k = \begin{pmatrix} -e_k \\ q_k \end{pmatrix}, \quad k = 1, \ldots, n \]
With CGS you could have catastrophic cancellation.

Gander 1980 - Better to compute Cholesky factorization of $A^T A$ and then set $Q = AR^{-1}$

Smoktunowicz, Barlow, and Langou, 2006
If $A^T A$ is numerically nonsingular and “the Pythagorean version of CGS” is used then the loss of orthogonality is proportional to $\kappa^2$
Pythagorean version of CGS

Computation of diagonal entry $r_{kk}$ at step $k$.

- **CGS:** $r_{kk} = \|w_k\|_2$ where $w_k = a_k - \sum_{i=1}^{k-1} r_{ik} q_i$

- **CGSP:** If

  $$s_k = \|a_k\|\quad \text{and} \quad p_k = \left( \sum_{i=1}^{k-1} r_{ik}^2 \right)^{1/2}$$

  then $r_{kk}^2 + p_k^2 = s_k^2$ and hence

  $$r_{kk} = (s_k - p_k)^{1/2}(s_k + p_k)^{1/2}$$
Heinz Rutishauser (1918–1970)

- Pioneer in Computer Science and Computational Mathematics
- Long history with ETH as both student and distinguished faculty
- Selective Reorthogonalization for MGS
- Superorthogonalization

Steven Leon, Åke Björck, Walter Gander
Gram-Schmidt Orthogonalization: 100 Years and More
Classic Gram-Schmidt

CGS Algorithm

\[ Q = A; \]
\[ \text{for } k = 1 : n \]
\[ \quad \text{for } i = 1 : k - 1; \]
\[ \quad \quad R(i, k) = Q(:, i)' \ast Q(:, k); \]
\[ \quad \text{end} \]
\[ \text{(Omit this line for CMGS)} \]
\[ \text{for } i = 1 : k - 1, \]
\[ \quad Q(:, k) = Q(:, k) - R(i, k) \ast Q(:, i); \]
\[ \text{(Omit this line for CMGS)} \]
\[ \text{end} \]
\[ R(k, k) = \text{norm}(Q(:, k)); \]
\[ Q(:, k) = Q(:, k) / R(k, k); \]
\[ \text{end} \]
CMGS Algorithm

\[ Q = A; \]

for \( k = 1 : n \)

\[ \text{for } i = 1 : k - 1 \]

\[ R(i, k) = Q(:, i)' \ast Q(:, k); \]

\[ Q(:, k) = Q(:, k) - R(i, k) \ast Q(:, i); \]

end

\[ R(k, k) = \text{norm}(Q(:, k)); \]

\[ Q(:, k) = Q(:, k)/R(k, k); \]

end
Iterative orthogonalization - As each $q_k$ is generated reorthogonalize with respect to $q_1, \ldots, q_{k-1}$

Twice is enough (W. Kahan), (Giraud, Langou, and Rozložník, 2002)

The algorithms MGS2, CGS2

Selective reorthogonalization
For most applications it is not necessary to reorthogonalize if MGS is used.

Both MGS least squares and MGS GMRES are backward stable.

In many applications it is important that the residual vector $r$ be orthogonal to the column vectors of $Q$.

In these cases if MGS is used, $r$ should be reorthogonalized with respect to $q_n, q_{n-1}, \ldots, q_1$ (note the reverse order).
Selective reorthogonalization

- Reorthogonalize when loss of orthogonality is detected.

\[ \mathbf{b} = \mathbf{a}_k - \sum_{i=1}^{k-1} r_{ik} \mathbf{q}_i \]

Indication of cancellation if \( \| \mathbf{b} \| \ll \| \mathbf{a}_k \| \)

- (1967) Rutishauser test: Reorthogonalize if \( \| \mathbf{b} \| \leq \frac{1}{10} \| \mathbf{a}_k \| \)
Selective reorthogonalization

- The Rutishauser reorthogonalization condition can be rewritten as

\[ \frac{\|a_k\|}{\|b\|} \geq 10 \text{ or more generally } \frac{\|a_k\|}{\|b\|} \geq K \]

- There are many papers using different values of \( K \).
  Generally \( 1 \leq K \leq \kappa_2(A) \); popular choice is \( K = \sqrt{2} \).
  Hoffman, 1989, investigates a range of \( K \) values.

- Giraud and Langou, 2003
  - Loss of orthogonality possible for any choice of \( K \)
  - Alternate condition

\[ \frac{\sum_{i=1}^{k-1} |r_{ik}|}{r_{kk}} > L \]
The vectors $\mathbf{x}$ and $\mathbf{y}$ are *numerically orthogonal* if

$$|f\ell(\mathbf{x}^T \mathbf{y})| \leq \epsilon \|\mathbf{x}\| \|\mathbf{y}\|$$  \hspace{1cm} (1)

In general

$$|\mathbf{x}^T \mathbf{y} - f\ell(\mathbf{x}^T \mathbf{y})| \leq \gamma_n |\mathbf{x}|^T |\mathbf{y}|$$  \hspace{1cm} (2)

$$\gamma_n = n\epsilon/(1 - n\epsilon), \ |\mathbf{x}| \text{ is the vector with elements } |x_i|$$

If $\mathbf{x}$ and $\mathbf{y}$ are orthogonal, then (2) becomes

$$|f\ell(\mathbf{x}^T \mathbf{y})| \leq \gamma_n |\mathbf{x}|^T |\mathbf{y}|$$  \hspace{1cm} (3)

We say that $\mathbf{x}$ and $\mathbf{y}$ are *numerically superorthogonal* if (3) is satisfied.
Rutishauser Orthno algorithm contains a routine for superorthogonalizing vectors.

The basic idea is to keep orthogonalizing \( \mathbf{x} \) and \( \mathbf{y} \) until (3) is satisfied.

The actual stopping condition used is:
While

\[
|\mathbf{x}|^T|\mathbf{y}| + \frac{|\mathbf{x}^T\mathbf{y}|}{10} > |\mathbf{x}|^T|\mathbf{y}|
\]

reorthogonalize \( \mathbf{x} \) and \( \mathbf{y} \)
Superorthogonalization Example

\[
\begin{pmatrix}
1 \\
10^{-40} \\
10^{-20} \\
10^{-10} \\
10^{-15}
\end{pmatrix}, \quad
\begin{pmatrix}
10^{-20} \\
1 \\
10^{-10} \\
10^{-20} \\
10^{-10}
\end{pmatrix}
\]

are numerically orthogonal, however

\[
|x^T y| \approx 10^{-20} \text{ and } |x|^T |y| \approx 10^{-20}
\]

Perform two reorthogonalizations: \( y \rightarrow y^{(2)} \rightarrow y^{(3)} \)
### Superorthogonalization Calculations

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
<th>( y^{(2)} )</th>
<th>( y^{(3)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1e+00</td>
<td>1e−20</td>
<td>−1.000019999996365e−25</td>
<td>−1.0000020000000000000e−25</td>
</tr>
<tr>
<td>1e−40</td>
<td>1e+00</td>
<td>1.00000000000000000000e+00</td>
<td>1.00000000000000000000e+00</td>
</tr>
<tr>
<td>1e−20</td>
<td>1e−10</td>
<td>1.00000000000000000000e−10</td>
<td>1.00000000000000000000e−10</td>
</tr>
<tr>
<td>1e−10</td>
<td>1e−20</td>
<td>9.99999998999989e−21</td>
<td>9.99999998999989e−21</td>
</tr>
<tr>
<td>1e−15</td>
<td>1e−10</td>
<td>1.00000000000000000000e−10</td>
<td>1.00000000000000000000e−10</td>
</tr>
</tbody>
</table>

Decrease in scalar products

\[
|x^T y| = 1e−20
\]
\[
|x^T y^{(2)}| = 3.6351e−37
\]
\[
|x^T y^{(3)}| = 0
\]
Conclusions

- Orthogonality plays a fundamental role in applied mathematics.
- The Gram-Schmidt algorithms are at the core of much of what we do in computational mathematics.
- Stability of GS is now well understood.
- The GS Process is central to solving least squares problems and to Krylov subspace methods.
- The QR factorization paved the way for modern rank revealing factorizations.
- What will the future bring?