Krylov Methods

Stefano Berrone

Politecnico di Torino, Dipartimento di Scienze Matematiche
stefano.berrone@polito.it

Computational Linear Algebra for Large Scale Problems
Data Science and Engineering
Politecnico di Torino

Last update: November 17, 2023

Introduction

Krylov subspace methods are **orthogonal or oblique projection methods onto Krylov subspaces**, which are subspaces spanned by vectors of the form $p(\mathrm{A})v$, where p is a polynomial.
In short, these techniques approximate the solution $\bar{x} = \mathrm{A}^{-1}b$ of the linear system $\mathrm{A}x = b$ by $p(\mathrm{A})b$.
A general projection method for solving the linear system $\mathrm{A}x = b$ is a method which seeks an approximate solution x_m from an affine space $x_0 + \mathcal{K}_m$ of dimension m by imposing the Galerkin/Petrov-Galerkin condition

$$b - \mathrm{A}x_m \perp \mathcal{L}_m,$$

where $\mathcal{L}_m$ is another subspace of dimension m, and x_0 is an **arbitrary initial guess**.

Given the initial residual $r_0 = b - \mathsf{A}x_0$, a Krylov subspace method is a method for which the subspace $\mathcal{K}_m$ is the Krylov subspace

$$\mathcal{K}_m(\mathsf{A}, r_0) = \mathrm{span}(r_0, \mathsf{A}r_0, \ldots, \mathsf{A}^{m-1}r_0).$$

Different Krylov methods arise from different choices of the test space $\mathcal{L}_m$ and are of the form

$$\bar{x} = \mathsf{A}^{-1}b \quad \approx \quad x_m = x_0 + q_{m-1}(\mathsf{A})r_0.$$

Krylov subspaces

Let us start introducing some definitions and properties of Krylov subspaces.
Given an arbitrary non-vanishing vector $v \in \mathbb{R}^n$, let us set

$$\mathcal{K}_m(\mathrm{A}, v) = \mathrm{span}(v, \mathrm{A}v, \dots, \mathrm{A}^{m-1}v).$$

If $\mathrm{A}^m v$ is linearly independent of the othe vectors $\mathrm{A}^j v$, $j = 0, \dots, m-1$ the dimension of the subspace of approximants increases by one at each step of the approximation process.

(Polynomial representation of Krylov subspace)

If $\{v, \mathrm{A}v, \dots, \mathrm{A}^{m-1}v\}$ are linearly independent, $\mathcal{K}_m(\mathrm{A}, v)$ is the subspace of $\mathbb{R}^n$ of dimension m whose elements can be written as $x = p(\mathrm{A})v$, where p is a polynomial of degree not exceeding $m-1$.

Definition (Minimal polyomial of A)

The **minimal polynomial of** A *is the nonzero monic polynomial* $p(x) = 1x^m + c_{m-1}x^{m-1} + c_1 x + c_0$ *of lowest degree such that* $p(\mathrm{A}) = 0$.

Theorem (Cayley-Hamilton)

Let A *be an* $n \times n$ *matrix, and let*
$p_{\mathrm{A}}(\lambda) = 1 * \lambda^n + c_{n-1}\lambda^{n-1} + \ldots + c_1\lambda + c_0$ *be its characteristic polynomial.*
Then $p_{\mathrm{A}}(\mathrm{A}) = 0$.

The Cayley-Hamilton theorem (A satisfies its characteristic equation) tell us that the minimal polynomial never exceeds n.

Property

If A *is diagonalizable the degree of the* **minimal polynomial of** A *is the number of distinct eigenvalues.*

Proof. Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ the distinct eigenvalues of A:

$$\begin{aligned} p(\lambda) &= (\lambda - \lambda_m)(\lambda - \lambda_{m-1}) \cdots (\lambda - \lambda_1) \\ &\quad = \lambda^m + a_{m-1}\lambda^{m-1} + a_1\lambda + a_0, \quad \text{(monic polynomial)} \\ p(\Lambda) &= (\Lambda - \lambda_m \mathrm{I})(\Lambda - \lambda_{m-1}\mathrm{I}) \cdots (\Lambda - \lambda_1 \mathrm{I}) = 0. \end{aligned}$$

Example: let us suppose that the matrix $A \in \mathbb{R}^{4,4}$ *has 3 distinct eigenvalues, with* λ_2 *with multiciplity 2.*

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{bmatrix}, \quad \Lambda - \lambda_1 I = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \lambda_2 - \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 - \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_3 - \lambda_1 \end{bmatrix},$$

$$\Lambda - \lambda_2 I = \begin{bmatrix} \lambda_1 - \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_3 - \lambda_2 \end{bmatrix}, \quad \Lambda - \lambda_3 I = \begin{bmatrix} \lambda_1 - \lambda_3 & 0 & 0 & 0 \\ 0 & \lambda_2 - \lambda_3 & 0 & 0 \\ 0 & 0 & \lambda_2 - \lambda_3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$(\Lambda - \lambda_1 I)(\Lambda - \lambda_2 I) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2) \end{bmatrix}.$$

The product $(\Lambda - \lambda_1 I)(\Lambda - \lambda_2 I)(\Lambda - \lambda_3 I)$ *trivially gives the zero matrix.*

Writing $A = P\Lambda P^{-1}$ we have

$$\begin{aligned} 0 = p(\Lambda) = {} & (\Lambda - \lambda_m I)(\Lambda - \lambda_{m-1} I) \cdots (\Lambda - \lambda_1 I) \\ P p(\Lambda) P^{-1} = {} & P(\Lambda - \lambda_m I)P^{-1} P(\Lambda - \lambda_{m-1} I)P^{-1} \cdots P(\Lambda - \lambda_1 I)P^{-1} \\ = {} & (A - \lambda_m I)(A - \lambda_{m-1} I) \cdots (A - \lambda_1 I) = p(A) = 0. \end{aligned}$$

Corollary (Cayley-Hamilton)

Let A *be an* $n \times n$ *matrix, and* $p(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \ldots + c_1\lambda + c_0$ *its characteristic polynomial,* A *is invertible if and only if* $c_0 \neq 0$*, and*

$$\mathrm{A}^{-1} = -\frac{1}{c_0}\left(\mathrm{A}^{n-1} + c_{n-1}\mathrm{A}^{n-2} + \ldots + c_1\mathrm{I}\right) = q_{\mathrm{inv}}(\mathrm{A}).$$

$$\mathrm{A}^n + c_{n-1}\mathrm{A}^{n-1} + \ldots + c_1\mathrm{A} + c_0\mathrm{I} = 0$$

$$\mathrm{A}\left(\mathrm{A}^{n-1} + c_{n-1}\mathrm{A}^{n-2} + \ldots + c_1\mathrm{I}\right) = -c_0\mathrm{I}$$

If A is invertible (no eigenvalue equal to zero), A^{-1} is a linear combination of $\mathrm{I}, \mathrm{A}, \ldots, \mathrm{A}^{m-1}$.
Note that
$c_0 = 0 \Leftrightarrow p(\lambda) = \lambda(\lambda^{n-1} + c_{n-1}\lambda^{n-2} + \ldots + c_1) \Leftrightarrow \lambda_n = 0.$

This property tells us that $\bar{x} = \mathrm{A}^{-1}b$ can be approximated by a linear combination of basis vectors of the Krylov space, in particular if $x_0 = 0$, $r_0 = b$ and $\mathcal{K}_m(\mathrm{A}, r_0) = \mathcal{K}_m(\mathrm{A}, b) = \mathrm{span}(b, \mathrm{A}b, \ldots, \mathrm{A}^{m-1}b)$ and $\bar{x}$ can be approximated by $x_m = q_{m-1}(\mathrm{A})b$.
If the chosen Krylov space is sufficiently large we can exactly represent $\bar{x} = \mathrm{A}^{-1}b = q_{\mathrm{inv}}(\mathrm{A})b$.

Definition (Minimal polyomial of a vector v with respect to A)

The **minimal polynomial of a vector** v **with respect to** A *is the nonzero monic polynomial* $p(x) = x^{\mu} + \ldots + c_{\mu-1}x^{\mu-1} + c_1 x + c_0$ *of lowest degree such that* $p_{\mu}(\mathrm{A})v = 0$*. The degree of the minimal polynomial is also called the* **grade of v with respect to** A*.*

Note that
$p_{\mu}(\mathrm{A})v = 0 \Leftrightarrow \{v, \mathrm{A}v, \mathrm{A}^2v, \ldots, \mathrm{A}^{\mu}v\}$ *not linearly independent*,
i.e. $\mathrm{A}^{\mu}v$ can be written as a linear combination of the vectors $\{v, \mathrm{A}v, \mathrm{A}^2v, \ldots, \mathrm{A}^{\mu-1}v\}$:
$\mathrm{A}^{\mu}v = -c_{\mu-1}\mathrm{A}^{\mu-1}v - \ldots - c_1\mathrm{A}v - c_0 v = -\sum_{j=0}^{\mu-1} c_j \mathrm{A}^j v$.

Proposition

Let μ be the grade of v with respect to A. Then $\mathcal{K}_\mu(\mathrm{A}, v)$ is invariant under A and $\mathcal{K}_m(\mathrm{A}, v) = \mathcal{K}_\mu(\mathrm{A}, v)$ for all $m \geq \mu$.

The dimension of the space $\mathcal{K}_m(\mathrm{A}, v)$ can not be arbitrarely increased increasing m, i.e. the vector $\mathrm{A}^m v$ for $m > \mu$ is linearly dependent on the vectors spanning $\mathcal{K}_\mu(\mathrm{A}, v)$.

Let us prove that if $\mathrm{A}^m v \in \mathcal{K}_\mu(\mathrm{A}, v)$ with $m > \mu$, then also $\mathrm{A}^{m+1} v \in \mathcal{K}_\mu(\mathrm{A}, v)$.

$$\mathrm{A}^\mu v = -\sum_{j=0}^{\mu-1} c_j \mathrm{A}^j v \qquad \mathrm{A}^m v = \sum_{j=0}^{\mu-1} \alpha_j \mathrm{A}^j v,$$

$$\begin{aligned}
\mathrm{A}\mathrm{A}^m v &= \mathrm{A}\sum_{j=0}^{\mu-1} \alpha_j \mathrm{A}^j v = \sum_{j=1}^{\mu-1} \alpha_{j-1} \mathrm{A}^j v + \alpha_{\mu-1} \mathrm{A}^\mu v \\
&= \sum_{j=1}^{\mu-1} \alpha_{j-1} \mathrm{A}^j v - \alpha_{\mu-1} \sum_{j=0}^{\mu-1} c_j \mathrm{A}^j v \\
&= -\alpha_{m-1} c_0 v + \sum_{j=1}^{\mu-1} (\alpha_{j-1} - \alpha_{\mu-1} c_j) \mathrm{A}^j v \in \mathcal{K}_\mu(\mathrm{A}, v).
\end{aligned}$$

The following proposition determines the dimension of $\mathcal{K}_m(\mathrm{A}, v)$ in general.

Proposition

The Krylov subspace $\mathcal{K}_m(\mathrm{A}, v)$ has dimension m if and only if the grade μ of v with respect to A is not less than m, i.e.

$$\dim(\mathcal{K}_m(\mathrm{A}, v)) = m \iff \mathrm{grade}(v, \mathrm{A}) \geq m.$$

Proof. ($\Rightarrow$) The vectors $v, \mathrm{A}v, \ldots, \mathrm{A}^{m-1}v$ form a basis of $\mathcal{K}_m(\mathrm{A}, v)$ (they are a set of generators), i.e. they are m vectors linearly independent $\Rightarrow$ the vector $p(\mathrm{A})v = 1\mathrm{A}^{m-1}v + \sum_{i=0}^{m-2} \alpha_i \mathrm{A}^i v \neq 0$ for any set of coefficients α_i, that means that $\mu > m - 1$.
($\Leftarrow$) If $\mathrm{grade}(v, \mathrm{A}) \geq m$, then any "monic polynomial" of degree $m - 1$ is non-zero $p(\mathrm{A})v \neq 0$, i.e. the vectors $v, \mathrm{A}v, \ldots, \mathrm{A}^{m-1}v$ are linearly independent.

Therefore, $\boxed{\dim(\mathcal{K}_m(\mathrm{A}, v)) = \min(m, \mathrm{grade}(v, \mathrm{A}))}$.

Basic Arnoldi Algorithm

Arnoldi's procedure is an algorithm for building an orthonormal basis of the Krylov subspaces $\mathcal{K}_m(\mathrm{A}, v_1)$.

Data: A, v_1 arbitrary vector of norm 1

Result: basis vectors of $\mathcal{K}_m(\mathrm{A}, v_1)$ in V, corresponding coefficients of $\mathrm{A}^j v_1$ in H

for $j = 1, \dots, m$ **do**
    for $i = 1, \dots, j$ **do**
        $h_{i,j} = (\mathrm{A}v_j, v_i)$;
    end
    $w_j = \mathrm{A}v_j - \sum_{i=1}^{j} h_{i,j} v_i$;
    $h_{j+1,j} = \|w_j\|_2$;
    if $h_{j+1,j} == 0$ **then return** $\mu = j$;
    $v_{j+1} = w_j / h_{j+1,j}$;
end

At each step the algorithm multiplies the previous Arnoldi vector v_j by A and then orthonormalizes the resulting vector $\mathrm{A}v_j$ against the previous v_i, $i = 1, \ldots, j$ by a standard Gram-Schmidt procedure.

Proposition

If the algorithm does not stop before the m-th step, the vectors $v_1, v_2, \ldots, v_m$ form an orthonormal basis of the Krylov subspace $\mathcal{K}_m(\mathrm{A}, v_1)$.

Proof. The vectors $v_1, v_2, \ldots, v_m$ are orthonormal by construction. The fact that they span $\mathcal{K}_m(\mathrm{A}, v_1)$ follows by induction from the fact that each v_j is of the form $q_{j-1}(\mathrm{A})v_1$, where q_{j-1} is a polynomial of degree $j-1$.

$$
\begin{aligned}
v_1 &= \frac{v_1}{\|v_1\|} = q_0(A)v_1, \\
v_2 &= \frac{Av_1 - (Av_1, v_1)v_1}{\|Av_1 - (Av_1, v_1)v_1\|} = q_1(A)v_1, \\
v_3 &= \frac{Av_2 - (Av_2, v_1)v_1 - (Av_2, v_2)v_2}{\|Av_2 - (Av_2, v_1)v_1 - (Av_2, v_2)v_2\|} \\
&= \frac{Aq_1(A)v_1 - (Av_2, v_1)v_1 - (Av_2, v_2)q_1(A)v_1}{\|Av_2 - (Av_2, v_1)v_1 - (Av_2, v_2)v_2\|} = q_2(A)v_1, \\
v_4 &= \frac{Av_3 - (Av_3, v_1)v_1 - (Av_3, v_2)v_2 - (Av_3, v_3)v_3}{\|Av_3 - (Av_3, v_1)v_1 - (Av_3, v_2)v_2 - (Av_3, v_3)v_3\|} \\
&= \frac{Aq_2(A)v_1 - (Av_3, v_1)v_1 - (Av_3, v_2)q_1(A)v_1 - (Av_3, v_3)q_2(A)v_1}{\|Av_3 - (Av_3, v_1)v_1 - (Av_3, v_2)v_2 - (Av_3, v_3)v_3\|} \\
&= q_3(A)v_1.
\end{aligned}
$$

Proposition

Denote by V_m *the* $n \times m$ *matrix with column vectors* $v_1, v_2, \ldots, v_m$, *by* $\bar{\mathrm{H}}_m$, *the* $(m+1) \times m$ *Hessemberg matrix whose nonzero entries* $h_{i,j}$ *are defined by the algorithm, and by* H_m *the matrix obtained by* $\bar{\mathrm{H}}_m$ *by deleting its last row. Then the following relations hold true:*

$$\begin{aligned} \mathrm{AV}_m &= \mathrm{V}_m\mathrm{H}_m + w_m e_m^T = \mathrm{V}_m\mathrm{H}_m + h_{m+1,m} v_{m+1} e_m^T, & (1)\\ &= \mathrm{V}_{m+1}\bar{\mathrm{H}}_m, & (2)\\ \mathrm{V}_m^T\mathrm{AV}_m &= \mathrm{H}_m. & (3) \end{aligned}$$

Proof. The previous relations derive from line 4 $w_j = \mathrm{A}v_j - \sum_{i=1}^{j} h_{i,j} v_i$ of the Arnoldi's algorithm that implies:

$$\mathrm{A}v_j = \sum_{i=1}^{j+1} h_{i,j} v_i = \sum_{i=1}^{j} h_{i,j} v_i + w_j, \qquad j = 1, 2, \ldots, m.$$

$e_m^T = [0, \ldots, 0, 1] \in \mathbb{R}^m$

$$
\begin{aligned}
Av_1 &= h_{1,1}v_1 + h_{2,1}v_2 \\
Av_2 &= h_{1,2}v_1 + h_{2,2}v_2 + h_{3,2}v_3 \\
Av_3 &= h_{1,3}v_1 + h_{2,3}v_2 + h_{3,3}v_3 + h_{4,3}v_4 \\
\ldots & \\
Av_m &= h_{1,m}v_1 + h_{2,m}v_2 + h_{3,m}v_3 + \ldots + h_{m,m}v_m + w_m \\
Av_m &= h_{1,m}v_1 + h_{2,m}v_2 + h_{3,m}v_3 + \ldots + h_{m,m}v_m + h_{m+1,m}v_{m+1}
\end{aligned}
$$

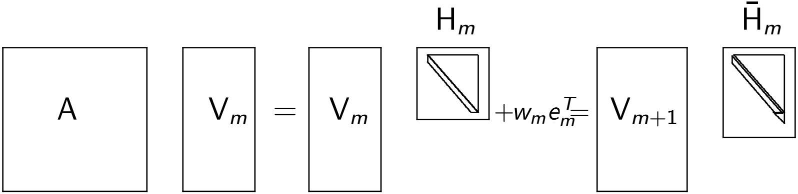

Figure: Action of A on V_m

Remark

$w_m e_m^T$ is a rank one $n \times m$ matrix with all zeros but in the m-th column where we have the column vector w_m.
Note that $V_{m+1}\bar{H}_m \in \mathbb{R}^{n\times m}$ also if the matrix $V_{m+1} \in \mathbb{R}^{n\times(m+1)}$

If the grade of v_1 with respect to A is $\mu = j$, $w_j = 0$ and any successive vector produced by the Arnoldi algorithm is zero.

Proposition

Arnoldi's algorithm breaks down at step j (i.e. $h_{j+1,j} = 0$), if and only if the minimal polynomial of v_1 is of degree j. Moreover, in this case the subspace $\mathcal{K}_j(\mathrm{A}, v_1)$ is invariant under A.

Proof. If the minimal polynomial of v_1 is of degree j, w_j must be equal to zero being $w_j = q_j(\mathrm{A})v_1$.

Modified Arnoldi Algorithm

Data: A, v_1 arbitrary vector of norm 1
Result: V, H

```
for j = 1, ..., m do
    w_j = A v_j;
    for i = 1, ..., j do
        h_{i,j} = (w_j, v_i);
        w_j = w_j - h_{i,j} v_i;
    end
    h_{j+1,j} = ||w_j||_2;
    if h_{j+1,j} == 0 then return μ = j;
    v_{j+1} = w_j / h_{j+1,j};
end
```

In exact arithmetic the Basic Arnoldi's Algorithm and the Modified Arnoldi's Algorithm are equivalent. In presence of round off the modified version is much more reliable. However, there are cases where cancellations are so severe in the orthogonalization steps that even the Modified Gram Schmidt option is inadequate. In these cases a double orthogonalization or an Huseholder orthogonalization technique can be useful.

Note that we can apply a QR factorization to the matrix V_{m+1} and use the relation $\mathrm{AV}_m = \mathrm{V}_{m+1}^{QR}\mathrm{R}\,\bar{\mathrm{H}}_m$ and define the improved matrix $\mathrm{V}_{m+1} = \mathrm{V}_{m+1}^{QR}$ and $\bar{\mathrm{H}}_m = \mathrm{R}\bar{\mathrm{H}}_m$.

The product of an upper Hessenberg matrix H with an upper triangular matrix U is upper Hessenberg: HU and UA are upper Hessenberg.

Given an initial guess x_0 ($r_0 = b - \mathrm{A}x_0$) we now consider an orthogonal projection method with

$$\mathcal{L} = \mathcal{K} = \mathcal{K}_m(\mathrm{A}, r_0).$$

This method seeks x_m from the affine subspace $x_0 + \mathcal{K}_m(\mathrm{A}, r_0)$ of dimension m by imposing the Galerkin ortogonality

$$r_m = b - \mathrm{A}x_m \perp \mathcal{K}_m(\mathrm{A}, r_0).$$

If $v_1 = r_0/\|r_0\|_2$, $\beta = \|r_0\|_2$, in the Arnoldi's method we have

$$\begin{aligned} \mathrm{V}_m^T \mathrm{A} \mathrm{V}_m &= \mathrm{H}_m, & r_m &= b - \mathrm{A}x_m \perp \mathcal{K}_m(\mathrm{A}, r_0) \\ \mathrm{V}_m^T r_0 = \mathrm{V}_m^T(\beta v_1) &= \beta e_1 & \mathrm{V}_m^T(b - \mathrm{A}(x_0 + \mathrm{V}_m y_m)) &= 0 \\ x_m &= x_0 + \mathrm{V}_m y_m & \mathrm{V}_m^T(r_0 - \mathrm{A}\mathrm{V}_m y_m) &= 0 \\ & & \beta e_1 - \mathrm{H}_m y_m &= 0 \end{aligned}$$

The resulting approximate solution in the m-dimensional subspace is given by

$$\mathrm{H}_m y_m = \beta e_1, \ \left(y_m = \mathrm{H}_m^{-1}(\beta e_1)\right).$$

FOM Full Orthogonalization Method

Data: A, b, x_0 arbitrary vector
Result: x_m

$r_0 = b - \mathrm{A}x_0$, $\beta = \|r_0\|_2$, $v_1 = r_0/\beta$;
for $j = 1, \ldots, m$ **do**
    $w_j = \mathrm{A}v_j$;
    for $i = 1, \ldots, j$ **do**
        $h_{i,j} = (w_j, v_i)$;
        $w_j = w_j - h_{i,j}v_i$;
    end
    $h_{j+1,j} = \|w_j\|_2$;
    if $h_{j+1,j} == 0$ **then break**;
    $v_{j+1} = w_j/h_{j+1,j}$;
end
$\mathrm{H}_m y_m = \beta e_1$, $x_m = x_0 + \mathrm{V}_m y_m$.

The above algorithm depends on the parameter dimension of the Krylov subspace m . The `for` loop at line 2 can be replaced by a do loop stopped when a given residual-based stopping criterion is satisfied. The target is to stop the iterations when $m \ll n$

Proposition

The residual vector of the approximate solution x_m computed by FOM is such that

$$r_m = b - \mathrm{A}x_m = -h_{m+1,m} v_{m+1} e_m^T y_m, \Rightarrow \|r_m\| = h_{m+1,m} |e_m^T y_m|.$$

Proof.

$$\begin{aligned} b - \mathrm{A}x_m &= b - \mathrm{A}(x_0 + \mathrm{V}_m y_m) = r_0 - \mathrm{A}\mathrm{V}_m y_m = r_0 - (\mathrm{V}_m \mathrm{H}_m + w_m e_m^T) y_m \\ &= \beta v_1 - \mathrm{V}_m \mathrm{H}_m y_m - h_{m+1,m} v_{m+1} e_m^T y_m = -h_{m+1,m} v_{m+1} e_m^T y_m \\ &= -h_{m+1,m} y_m(m) v_{m+1}. \end{aligned}$$

By definition of y_m: $\mathrm{H}_m y_m = \beta e_1$ and then $\beta v_1 - \mathrm{V}_m \mathrm{H}_m y_m = \beta v_1 - \mathrm{V}_m \beta e_1 = 0.$

FOM Full Orthogonalization Method

Data: A, b, x_0 arbitrary vector
Result: x_m

$m = 0,\ r_0 = b - \mathrm{A}x_0,\ \beta = \|r_0\|_2,\ v_1 = r_0/\beta$;
repeat
    $m = m + 1,\ V(:, m) = v_m,\ w_m = \mathrm{A}v_m$;
    if $m > 1$ $H(m, m-1) = \gamma$;
    for $i = 1, \ldots, m$ **do**
        $H(i, m) = (w_m, v_i)$;
        $w_m = w_m - H(i, m)v_i$;
    end
    $\gamma = \|w_m\|_2,\ v_{m+1} = w_m/\gamma$;
    if $m > 1$ **Compute** P, $\mathrm{PH}y = \beta \mathrm{P}e_1$ **s.t.** $H(m, m-1) = 0$;
    Compute $y(m)$
until $|\gamma y(m)| < tol$;
$\mathrm{H}y = \beta e_1,\ x_m = x_0 + \mathrm{V}y$.

Restarted FOM

As m increases the computational cost increases at least as $\mathcal{O}(m^2 n)$ because the Gram Schmidt ortoghonalization. The memory cost increases as $\mathcal{O}(mn)$. For very large n this limits the largest values of m that can be used. A possible remedy is to restart the algorithm periodically.

Data: A, b, x_0 arbitrary vector, M restart dimension
Result: x_m

repeat
    $r_0 = b - \mathrm{A}x_0$, $\beta = \|r_0\|_2$, $v_1 = r_0/\beta$;
    Compute H_M and V_M using Arnoldi algorithm;
    $y_m = \mathrm{H}_M^{-1}(\beta e_1)$, $x_m = x_0 + \mathrm{V}_M y_M$;
    $x_0 = x_m$;
until stopping criterion;

Generalized Minimal Residual Method (GMRES)

GMRES is a projection method based on taking $\mathcal{K} = \mathcal{K}_m(\mathrm{A}, r_0)$ and $\mathcal{L} = \mathrm{A}\mathcal{K}_m(\mathrm{A}, r_0)$ and looking for $r_m = b - \mathrm{A}x_m \perp \mathcal{L}$.
Any vector $x_m \in x_0 + \mathcal{K}_m(\mathrm{A}, r_0)$ can be written as

$$x_m = x_0 + \mathrm{V}_m y, \qquad y \in \mathbb{R}^m.$$

$$\begin{aligned} r_m = b - \mathrm{A}x_m &= b - \mathrm{A}(x_0 + \mathrm{V}_m y) = r_0 - \mathrm{V}_{m+1}\bar{\mathrm{H}}_m y \\ &= \beta v_1 - \mathrm{V}_{m+1}\bar{\mathrm{H}}_m y = \mathrm{V}_{m+1}(\beta e_1 - \bar{\mathrm{H}}_m y) \end{aligned}$$

Let us impose the orthogonality condition with respect to $\mathcal{L} = \mathrm{A}\mathcal{K}_m(\mathrm{A}, r_0)$:

$$\begin{aligned} (\mathrm{A}\mathrm{V}_m)^T(\mathrm{V}_{m+1}(\beta e_1 - \bar{\mathrm{H}}_m y)) &= \bar{\mathrm{H}}_m^T \mathrm{V}_{m+1}^T \mathrm{V}_{m+1}(\beta e_1 - \bar{\mathrm{H}}_m y) = 0 \\ \bar{\mathrm{H}}_m^T \bar{\mathrm{H}}_m y &= \bar{\mathrm{H}}_m^T \beta e_1. \end{aligned}$$

This condition is exactly the condition that we get by choosing x_m as the minimizer of $\|r_m\|_2^2 = \|b - \mathrm{A}x_m\|_2^2$, in fact

$$
\begin{aligned}
\|r_m\|_2^2 &= (\mathrm{V}_{m+1}(\beta e_1 - \bar{\mathrm{H}}_m y))^T \mathrm{V}_{m+1}(\beta e_1 - \bar{\mathrm{H}}_m y) \\
&= \beta^2 - 2\beta y^T \bar{\mathrm{H}}_m^T e_1 + y^T \bar{\mathrm{H}}_m^T \bar{\mathrm{H}}_m y. \\
\nabla \|r_m\|_2^2 &= -2\beta \bar{\mathrm{H}}_m^T e_1 + 2\bar{\mathrm{H}}_m^T \bar{\mathrm{H}}_m y. \\
\nabla \|r_m\|_2^2 = 0 &\Leftrightarrow \bar{\mathrm{H}}_m^T \bar{\mathrm{H}}_m y = \bar{\mathrm{H}}_m^T \beta e_1.
\end{aligned}
$$

Moreover, let us note that $\min \|b - \mathrm{A}x_m\|_2^2 \Leftrightarrow \min \|b - \mathrm{A}x_m\|_2$. Instead of solving the "normal equation" system $\bar{\mathrm{H}}_m^T \bar{\mathrm{H}}_m y = \bar{\mathrm{H}}_m^T \beta e_1$ we prefer to minimize the functional $J(y) = \|b - \mathrm{A}x_m\|_2$.

Since the column-vectors of V_{m+1} are orthonormal, then

$$\begin{aligned}
\|x\|_2^2 &= x^T x = x^T \mathrm{V}_{m+1}^T \mathrm{V}_{m+1} x = \|\mathrm{V}_{m+1} x\|_2^2, \\
J(y) &= \|b - \mathrm{A}x_m\|_2 = \|b - \mathrm{A}(x_0 + \mathrm{V}_m y)\|_2 \\
&= \|\beta \mathrm{V}_{m+1} e_1 - \mathrm{V}_{m+1} \bar{\mathrm{H}}_m y\|_2 = \|\beta e_1 - \bar{\mathrm{H}}_m y\|_2.
\end{aligned}$$

The GMRES approximation is the unique vector $x_0 + \mathcal{K}_m(\mathrm{A}, r_0)$ which minimizes $J(x) = \|b - \mathrm{A}x\|_2$. The approximation $x_m = x_0 + \mathrm{V}_m y_m$ ($r_m = r_0 - \mathrm{A}\mathrm{V}_m y_m = r_0 - \mathrm{V}_{m+1} \bar{\mathrm{H}}_m y_m$) is defined by

$$y_m = \underset{y \in \mathbb{R}^m}{\operatorname{argmin}} \|\beta e_1 - \bar{\mathrm{H}}_m y\|_2.$$

The minimizer y_m requires the solution of an $(m+1) \times m$ least-square problem. The minimizer y_m is the vector of coefficients with respect to the basis V_m of the correction $x_m - x_0$. Moreover, the vector $\bar{\mathrm{H}}_m y_m$ describes the correction $r_m - r_0$ with respect to V_{m+1}.

GMRES

Data: A, b, x_0 arbitrary vector

Result: x_m

$r_0 = b - \mathrm{A}x_0$, $\beta = \|r_0\|_2$, $v_1 = r_0/\beta$;

for $j = 1, \dots, m$ **do**

    $w_j = \mathrm{A}v_j$;

    for $i = 1, \dots, j$ **do**

        $h_{i,j} = (w_j, v_i)$;

        $w_j = w_j - h_{i,j}v_i$;

    end

    $h_{j+1,j} = \|w_j\|_2$;

    if $h_{j+1,j} == 0$ **then break**;

    $v_{j+1} = w_j / h_{j+1,j}$;

end

$y_m =$ minimizer of $\|\beta e_1 - \bar{\mathrm{H}}_m y\|_2$, $x_m = x_0 + \mathrm{V}_m y_m$.

The loop from line 2 to line 9 is the Arnoldi algorithm based on the modified Gram-Schmidt orthonormalization needed to generate the matrices V_m and $\bar{\mathrm{H}}_m$. A different approach based on the Householder orthogonalization is possible and preferable due an increased numerical stability. Some libraries allow to chose the orthogonalization algorithm for the GMRES method.
A final issue concerns the solution of the rectangular $(m+1)\times m$ least-square problem.

Proposition (part I)

Let

$$\begin{aligned} \mathrm{Q}_m^T \bar{\mathrm{R}}_m &= \bar{\mathrm{H}}_m, \\ \bar{g}_m &= \mathrm{Q}_m(\beta e_1) = [\gamma_1, \gamma_2, \dots, \gamma_{m+1}]^T = [g_m, \gamma_{m+1}]^T, \end{aligned}$$

where Q_m *is an* **orthogonal** $(m+1) \times (m+1)$ **matrix** *and* $\bar{\mathrm{R}}_m$ *is an* **upper triangular** $(m+1) \times m$ **matrix** *(all zeros in row* $m+1$*). Moreover,* R_m *is the* **upper triangular** $m \times m$ **part** *of* $\bar{\mathrm{R}}_m$ *obtained deleting the last row of* $\bar{\mathrm{R}}_m$ *that contains only zeros and* g_m *if the m-dimensional vector obtained from* $\bar{g}_m$ *by deleting the last component. Then,*

Note that the first m columns of the matrix Q_m are defined by the columns of $\bar{\mathrm{H}}_m$, whereas the last column is arbitrary, but orthogonal to the previous columns.

Proposition (part II)

1. *The rank of* AV_m *is equal to the rank of* R_m. *In particular, if* $r_{m,m} = 0$, *then* A *must be singular.*
2. *The vector* y_m *which minimizes* $\|\beta e_1 - \bar{\mathrm{H}}_m y\|$ *is given by*
$$y_m = \mathrm{R}_m^{-1} g_m.$$
3. *The residual vector at step m satisfies*
$$b - \mathrm{A}x_m = \mathrm{V}_{m+1}(\beta e_1 - \bar{\mathrm{H}}_m y_m) = \mathrm{V}_{m+1}\mathrm{Q}_m^T(\gamma_{m+1}e_{m+1})$$
and, as result,
$$\|b - \mathrm{A}x_m\|^2 = \|\beta e_1 - \bar{\mathrm{H}}_m y_m\|^2 = \|\mathrm{Q}_m(\beta e_1 - \bar{\mathrm{H}}_m y_m)\|^2$$
$$= \|\bar{g}_m - \bar{\mathrm{R}}_m y\|^2 = \|g_m - \mathrm{R}_m y\|^2 + |\gamma_{m+1}|^2 = |\gamma_{m+1}|^2.$$

Proof.

1 $$\mathrm{A}\mathrm{V}_m = \mathrm{V}_{m+1}\bar{\mathrm{H}}_m = \mathrm{V}_{m+1}\mathrm{Q}_m^T\bar{\mathrm{R}}_m.$$

Since $\mathrm{V}_{m+1}\mathrm{Q}_m^T$ is $n \times (m+1)$ unitary:

$$(\mathrm{V}_{m+1}\mathrm{Q}_m^T)^T\mathrm{V}_{m+1}\mathrm{Q}_m^T = \mathrm{Q}_m\mathrm{V}_{m+1}^T\mathrm{V}_{m+1}\mathrm{Q}_m^T = \mathrm{Q}_m\mathrm{I}_{m+1}\mathrm{Q}_m^T = \mathrm{I}_{m+1},$$

the rank of $\mathrm{A}\mathrm{V}_m$ is that of $\bar{\mathrm{R}}_m$, which equals the rank of R_m since these two matrices differs only by a zero row (the last of $\bar{\mathrm{R}}_m$). If $r_{m,m} = 0$ then R_m is of $\mathrm{rank} \leq m-1$ and $\mathrm{A}\mathrm{V}_m$ is also of $\mathrm{rank} \leq m-1$. Since V_m is full rank, this means that A must be singular.

Proof.

2 For any vector y

$$
\begin{aligned}
\|\beta e_1 - \bar{\mathrm{H}}_m y\|_2^2 &= \|\mathrm{Q}_m(\beta e_1 - \bar{\mathrm{H}}_m y)\|_2^2 = \|\bar{g}_m - \bar{\mathrm{R}}_m y\|_2^2 \\
&= |\gamma_{m+1}|^2 + \|g_m - \mathrm{R}_m y\|_2^2.
\end{aligned}
$$

Since R_m is nonsingular, the minimum of the left hand side is reached for $y = \mathrm{R}_m^{-1} g_m$.

3

$$
\begin{aligned}
b - \mathrm{A}x &= \mathrm{V}_{m+1}(\beta e_1 - \bar{\mathrm{H}}_m y) = \mathrm{V}_{m+1}\mathrm{Q}_m^T \mathrm{Q}_m(\beta e_1 - \bar{\mathrm{H}}_m y) \\
&= \mathrm{V}_{m+1}\mathrm{Q}_m^T(\bar{g}_m - \bar{\mathrm{R}}_m y) = \mathrm{V}_{m+1}\mathrm{Q}_m^T(\gamma_{m+1} e_{m+1}).
\end{aligned}
$$

The norm $\|b - \mathrm{A}x_m\| = |\gamma_{m+1}|$ thanks to the orthonormality of the column vectors of $\mathrm{V}_{m+1}\mathrm{Q}_m^T$.

Givens plane rotation matrices

$$\mathrm{G} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}.$$

The matrix G performs a counterclockwise rotation of the reference system of an angle θ, or equivalently, when applied to a vector x, Gx is the same vector x rotated clockwise of an angle θ.
The Givens matrix G is an orthogonal matrix

$$\mathrm{G}^{-1} = \begin{pmatrix} \cos(-\theta) & \sin(-\theta) \\ -\sin(-\theta) & \cos(-\theta) \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = \mathrm{G}^T.$$

Given a vector h the rotation of our interest is the rotation that transforms x to a vector with a vanishing second component.

$$\mathrm{G}h = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix},$$

where the angle θ is such that

$$\cos(\theta) = \frac{h_1}{\sqrt{h_1^2 + h_2^2}}, \qquad \sin(\theta) = \frac{h_2}{\sqrt{h_1^2 + h_2^2}}.$$

The generic Givens matrix is

$$\Omega_i = G_{i,i+1} = \begin{array}{c c} & \begin{array}{cccccccc} & & & & i & i+1 & & \end{array} \\ \begin{array}{c} \\ \\ \\ \\ i \\ i+1 \\ \\ \\ \end{array} & \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ & & & & \vdots & \vdots & & \\ 0 & 0 & 0 & \dots & \cos(\theta) & \sin(\theta) & \dots & 0 \\ 0 & 0 & 0 & \dots & -\sin(\theta) & \cos(\theta) & \dots & 0 \\ & & & & \vdots & \vdots & & \\ 0 & 0 & 0 & \dots & 0 & \dots & \dots & 1 \end{pmatrix} \end{array}.$$

This matrix can be applied to the matrix $\bar{H}_m$ in order to transform it to an upper triangular matrix iteratively vanishing the elements under the diagonal. The vanishing of each element under the diagonal requires the multiplication by a suitable Given matrix.

Let us consider the following 6×5 example

$$\bar{\mathrm{H}}_5 = \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} & h_{15} \\ h_{21} & h_{22} & h_{23} & h_{24} & h_{25} \\ 0 & h_{32} & h_{33} & h_{34} & h_{35} \\ 0 & 0 & h_{43} & h_{44} & h_{45} \\ 0 & 0 & 0 & h_{54} & h_{55} \\ 0 & 0 & 0 & 0 & h_{65} \end{pmatrix} \in \mathbb{R}^{6\times 5},$$

$$\bar{\Omega}_1 \bar{\mathrm{H}}_5 = \begin{pmatrix} c_1 & s_1 & 0 & 0 & 0 & 0 \\ -s_1 & c_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \bar{\mathrm{H}}_5 = \bar{\mathrm{H}}_5^{(1)} = \begin{pmatrix} h_{11}^{(1)} & h_{12}^{(1)} & h_{13}^{(1)} & h_{14}^{(1)} & h \\ 0 & h_{22}^{(1)} & h_{23}^{(1)} & h_{24}^{(1)} & h \\ 0 & h_{32}^{(1)} & h_{33}^{(1)} & h_{34}^{(1)} & h \\ 0 & 0 & h_{43}^{(1)} & h_{44}^{(1)} & h \\ 0 & 0 & 0 & h_{54}^{(1)} & h \\ 0 & 0 & 0 & 0 & h \end{pmatrix}$$

$$c_1 = \frac{h_{11}}{\sqrt{h_{11}^2 + h_{21}^2}}, \qquad s_1 = \frac{h_{21}}{\sqrt{h_{11}^2 + h_{21}^2}}$$

$$\bar{\Omega}_2\bar{\mathrm{H}}_5^{(1)}=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0\\ 0 & c_2 & s_2 & 0 & 0 & 0\\ 0 & -s_2 & c_2 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}\bar{\mathrm{H}}_5^{(1)}=\bar{\mathrm{H}}_5^{(2)}=\begin{pmatrix} h_{11}^{(2)} & h_{12}^{(2)} & h_{13}^{(2)} & h_{14}^{(2)} & h_{15}^{(2)}\\ 0 & h_{22}^{(2)} & h_{23}^{(2)} & h_{24}^{(2)} & h_{25}^{(2)}\\ 0 & 0 & h_{33}^{(2)} & h_{34}^{(2)} & h_{35}^{(2)}\\ 0 & 0 & h_{43}^{(2)} & h_{44}^{(2)} & h_{45}^{(2)}\\ 0 & 0 & 0 & h_{54}^{(2)} & h_{55}^{(2)}\\ 0 & 0 & 0 & 0 & h_{65}^{(2)} \end{pmatrix},$$

$$c_2 = \frac{h_{11}^{(1)}}{\sqrt{{h_{11}^{(1)}}^2 + {h_{21}^{(1)}}^2}}, \qquad s_2 = \frac{h_{21}^{(1)}}{\sqrt{{h_{11}^{(1)}}^2 + {h_{21}^{(1)}}^2}}.$$

After 5 steps we have

$$Q_5 = \Omega_5\Omega_4\Omega_3\Omega_2\Omega_1 \in \mathbb{R}^{6\times 6}.$$

$$\bar{H}_5^{(5)} = \begin{pmatrix} h_{11}^{(5)} & h_{12}^{(5)} & h_{13}^{(5)} & h_{14}^{(5)} & h_{15}^{(5)} \\ 0 & h_{22}^{(5)} & h_{23}^{(5)} & h_{24}^{(5)} & h_{25}^{(5)} \\ 0 & 0 & h_{33}^{(5)} & h_{34}^{(5)} & h_{35}^{(5)} \\ 0 & 0 & 0 & h_{44}^{(5)} & h_{45}^{(5)} \\ 0 & 0 & 0 & 0 & h_{55}^{(5)} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \bar{g}_5 = Q_5(\beta e_1) = \begin{pmatrix} g_5 = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \end{pmatrix} \\ \gamma_6 \end{pmatrix}.$$

In general

$$Q_m = \Omega_m\Omega_{m-1}\ldots\Omega_2\Omega_1 \in \mathbb{R}^{(m+1)\times(m+1)}.$$

It is possible to implement the factorization process in a progressive manner at each step of the GMRES algorithm. This approach will allow us to obtain the residual norm at each step of the algorithm, with essentially no additional arithmetic operations. In the previous example at step 6 we have

$$
\begin{pmatrix}
h_{11}^{(5)} & h_{12}^{(5)} & h_{13}^{(5)} & h_{14}^{(5)} & h_{15}^{(5)} & h_{16} \\
0 & h_{22}^{(5)} & h_{23}^{(5)} & h_{24}^{(5)} & h_{25}^{(5)} & h_{26} \\
0 & 0 & h_{33}^{(5)} & h_{34}^{(5)} & h_{35}^{(5)} & h_{36} \\
0 & 0 & 0 & h_{44}^{(5)} & h_{45}^{(5)} & h_{46} \\
0 & 0 & 0 & 0 & h_{55}^{(5)} & h_{56} \\
0 & 0 & 0 & 0 & 0 & h_{66} \\
0 & 0 & 0 & 0 & 0 & h_{76}
\end{pmatrix}
=
\begin{pmatrix}
 & & & & & h_{16} \\
0 & & & & & h_{26} \\
0 & 0 & & \mathrm{R}_5 & & h_{36} \\
0 & 0 & 0 & & & h_{46} \\
0 & 0 & 0 & 0 & & h_{56} \\
0 & 0 & 0 & 0 & 0 & h_{66} \\
0 & 0 & 0 & 0 & 0 & h_{76}
\end{pmatrix}
, \bar{g}_6^{(5)} =
\begin{pmatrix}
\gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \\ \gamma_6 \\ 0
\end{pmatrix}.
$$

Where we have appended to $\bar{\mathrm{H}}_5^{(5)} = \bar{\mathrm{R}}_5$ the sixth column of $\bar{\mathrm{H}}_6$ and filled the remaining part of the last row with zeros.

We apply to the **last column** the othogonal matrix Q_5 obtaining

$$\bar{\mathrm{H}}_6^{(5)} = \begin{pmatrix} h_{11}^{(5)} & h_{12}^{(5)} & h_{13}^{(5)} & h_{14}^{(5)} & h_{15}^{(5)} & h_{16}^{(5)} \\ 0 & h_{22}^{(5)} & h_{23}^{(5)} & h_{24}^{(5)} & h_{25}^{(5)} & h_{26}^{(5)} \\ 0 & 0 & h_{33}^{(5)} & h_{34}^{(5)} & h_{35}^{(5)} & h_{36}^{(5)} \\ 0 & 0 & 0 & h_{44}^{(5)} & h_{45}^{(5)} & h_{46}^{(5)} \\ 0 & 0 & 0 & 0 & h_{55}^{(5)} & h_{56}^{(5)} \\ 0 & 0 & 0 & 0 & 0 & h_{66}^{(5)} \\ 0 & 0 & 0 & 0 & 0 & h_{76} \end{pmatrix}.$$

Note that we can apply the action of the matrix Q_5 just saving the coefficients c_i, s_i, $i = 1, \ldots, 5$ considering that the matrices Ω_i affects just the rows i and $i+1$.

After this we compute Ω_6 in the usual way and we apply it to the full matrix $\bar{\mathrm{H}}_6^{(5)}$ obtaining

$$\bar{\mathrm{H}}_6^{(6)} = \begin{pmatrix} h_{11}^{(6)} & h_{12}^{(6)} & h_{13}^{(6)} & h_{14}^{(6)} & h_{15}^{(6)} & h_{16}^{(6)} \\ 0 & h_{22}^{(6)} & h_{23}^{(6)} & h_{24}^{(6)} & h_{25}^{(6)} & h_{26}^{(6)} \\ 0 & 0 & h_{33}^{(6)} & h_{34}^{(6)} & h_{35}^{(6)} & h_{36}^{(6)} \\ 0 & 0 & 0 & h_{44}^{(6)} & h_{45}^{(6)} & h_{46}^{(6)} \\ 0 & 0 & 0 & 0 & h_{55}^{(6)} & h_{56}^{(6)} \\ 0 & 0 & 0 & 0 & 0 & h_{66}^{(6)} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \bar{g}_6^{(6)} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \\ c_6\gamma_6 \\ -s_6\gamma_6 \end{pmatrix}.$$

From the previous relation we can see that the residual norm $|\gamma_{m+1}|$ is given by $|-s_m\gamma_m| = |(-1)^m\beta\prod_1^m s_m|$. If this quantity is small enough, the algorithm stops and then it uses the matrix R_m and the vector g_m to solve the upper triangular system $\mathrm{R}_m y_m = g_m$ to compute y_m and then the solution $x_m = x_0 + \mathrm{V}_m y_m$.

At each GMRES iteration we need to compute/update $\bar{\mathrm{H}}_m$, V_m, Q_m and $\bar{g}_m$. We can stop the GMRES iterations using the last component of $\bar{g}_m$ end we never need to compute the solution x_m or the residual r_m before stopping the loop.

Proposition

Let A be a nonsingular matrix, and let the grade of r0 with respect to A larger than j. Then, the GMRES algorithm breaks down at step j, i.e. $h_{j+1,j} = 0$, if and only if the approximate solution x_j is exact.

Proof. To show "$h_{j+1,j} = 0 \Rightarrow x_j$ is the exact solution", observe that if $h_{j+1,j} = 0$, then the matrix Ω_j has $s_j = h_{j+1,j} / \sqrt{\left(h_{j,j}^{(j-1)}\right)^2 + h_{j+1,j}^2} = 0$. Since A is nonsingular, then $r_{j,j} = h_{j,j}^{(j-1)}$ is nonzero by the first point of previous proposition (The rank of AV_m is equal to the rank of R_m. In particular, if $r_{m,m} = 0$, then A must be singular) and this implies that the denominator is nonzero and $s_j = 0$. Then, $|\gamma_{j+1}| = |-s_j\gamma_j| = \|b - \mathrm{A}x_j\|$.

Vice versa, if the solution x_j is exact and not x_{j-1}, then $\|b - \mathrm{A}x_j\| = |\gamma_{j+1}| = |-s_j\gamma_j| = 0$ and

$\|b - \mathrm{A}x_{j-1}\| = |\gamma_j| \neq 0$ and this implies $s_j = 0$, that implies $h_{j+1,j} = 0$.

Restarted GMRES

Similar to the FOM algorithm, the GMRES algorithm becomes impractical when m is large because of the growth of memory and computational requirements as m increase. A possible solution to this problem is to restart it after a given number of iterations.
A well known difficulty with the restarted GMRES algorithm is that it can *stagnate* when the matrix is not positive definite. The full GMRES algorithm is guaranteed to converge in at most n steps, but this would be impractical.

Elementary reflection matrices or Householder reflectors

Definition

An elementary reflection matrix (reflector) in $\mathbb{R}^n$ is a matrix of type

$$\mathrm{P} = \mathrm{I} - 2uu^T,$$

where u is a unit vector in $\mathbb{R}^n$.

Linear transformations associated o these matrices are called *Householder's transformations*.

The following properties can be easily verified:

- P is symmetric: $\mathrm{P}^T = \mathrm{P}$;
- P is orthogonal: $\mathrm{P}^{-1} = \mathrm{P}^T$;
- P is involutive: $\mathrm{P}^2 = \mathrm{I}$.

Let us consider the effect of the matrix P. Let us take a generic nonvanishing vector x and let v be its component along the vector u, $v = \alpha u$, $\alpha = u^T x$.

Moreover, let w be its component othogonal to u. The matrix P acts a reflection of the vector x with respect to the line orthogonal to the vector u (vectors applied in the origin of the reference system). Let us remark that $\mathrm{P}x - x = -2\alpha u$, this implies that the vector u is parallel to the vector $\mathrm{P}x - x$.

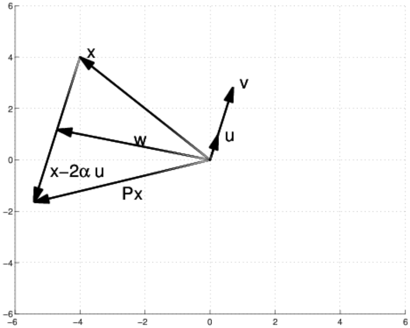

Figure: Example of Householder's reflection

Remark

Let x and y be two nonvanishing vectors in $\mathbb{R}^n$, linearly independent. A vector $u \in \mathbb{R}^n$ with $\|u\|_2 = 1$ and a scalar σ can be found such that

$$Px = (I - 2uu^T)x = \sigma y.$$

Indeed

$$\|Px\|_2 = \|x\|_2 = \|\sigma y\|_2 = |\sigma| \|y\|_2$$

and

$$|\sigma| = \frac{\|x\|_2}{\|y\|_2} \quad \Rightarrow \quad \sigma = \pm \frac{\|x\|_2}{\|y\|_2}.$$

Remark

We already observed that the unit vector u is parallel to $\mathrm{P}x - x$, so

$$u = \frac{x - \mathrm{P}x}{\|x - \mathrm{P}x\|_2} = \frac{x - \sigma y}{\|x - \sigma y\|_2}.$$

For each choice of the signum of σ we have a different vector u, for stability reasons, the preferable choice is that provides the largest vector $x - \mathrm{P}x = x - \sigma y$ whose norm is at the denominator in the definition of u.

Very often the Householder reflectors are applied in order to vanish the last components of a given vector.

Theorem

For each given nonvanishing vector x, the Householder's reflector

$$P = I - 2\frac{\tilde{u}\tilde{u}^T}{\|\tilde{u}\|_2^2},$$

with $\tilde{u} = x + \sigma e_1$ and $\sigma = \mathrm{sign}(x_1)\|x\|_2$, reflects the vector x on the direction e_1:

$$Px = -\sigma e_1.$$

Proof.

Indeed:

$$\mathrm{P}x = \left(\mathrm{I} - 2\frac{\tilde{u}\tilde{u}^T}{\|\tilde{u}\|_2^2}\right)x = x - 2\frac{\tilde{u}^T x}{\|\tilde{u}\|_2^2}\tilde{u}.$$

$$\begin{aligned}\tilde{u}^T x &= (x + \sigma e_1)^T x = \left(x^T + \sigma e_1^T\right)x = x^T x + \sigma e_1^T x \\ &= \sigma^2 + \sigma x_1 = \sigma(\sigma + x_1),\end{aligned}$$

$$\tilde{u}^T\tilde{u} = (x + \sigma e_1)^T(x + \sigma e_1) = x^T x + 2\sigma x_1 + \sigma^2 = 2\sigma(\sigma + x_1).$$

Concluding,

$$\mathrm{P}x = x - \tilde{u} = -\sigma e_1.$$

$\square$

Remark

In the choice of the signum of σ we have performed stability arguments.