- GMRES – Generalized Minimal Residual
- Conjugate Gradients.

GMRES – Generalized Minimal Residuals

- Arnoldi method can be used to find eigenvalues. GMRES is its analog for $Ax = b$ problems.

GMRES – Generalized Minimal Residuals

- Arnoldi method can be used to find eigenvalues. GMRES is its analog for $Ax = b$ problems.
- Let $\mathcal{K}_n$ denote the Krylov subspace $\langle b, Ab, \ldots, A^{n-1}b \rangle$ – then GMRES tries to do the following:

GMRES – Generalized Minimal Residuals

- Arnoldi method can be used to find eigenvalues. GMRES is its analog for $Ax = b$ problems.
- Let $\mathcal{K}_n$ denote the Krylov subspace $\langle b, Ab, \ldots, A^{n-1}b \rangle$ – then GMRES tries to do the following:
- At step n, one approximates the exact solution, x_* (ie $Ax_* = b$), by the vector $x_n \in \mathcal{K}_n$ that minimizes the norm of the residual, $r_n = b - Ax_n$.

GMRES – Generalized Minimal Residuals

- Arnoldi method can be used to find eigenvalues. GMRES is its analog for $Ax = b$ problems.
- Let $\mathcal{K}_n$ denote the Krylov subspace $\langle b, Ab, \ldots, A^{n-1}b \rangle$ – then GMRES tries to do the following:
- At step n, one approximates the exact solution, x_* (ie $Ax_* = b$), by the vector $x_n \in \mathcal{K}_n$ that minimizes the norm of the residual, $r_n = b - Ax_n$.
- In other words, the x_n is determined by solving a least squares problem.

- One obvious way to solve this problem would be the following.

- One obvious way to solve this problem would be the following.
- Let K_n be the Krylov matrix so that

$$AK_n = \left[\begin{array}{c|c|c|c} & & & \\ Ab & A^2b & \cdots & A^nb \\ & & & \end{array} \right]$$

- One obvious way to solve this problem would be the following.
- Let K_n be the Krylov matrix so that

$$AK_n = \left[\begin{array}{c|c|c|c} & & & \\ Ab & A^2b & \cdots & A^nb \\ & & & \end{array} \right]$$

- The column space of this matrix is $A\mathcal{K}_n$. Thus the problem reduces to finding a vector $c \in \mathbb{C}^n$ such that

$$||AK_nc - b|| = minimum$$

GMRES – Generalized Minimal Residuals

- One obvious way to solve this problem would be the following.
- Let K_n be the Krylov matrix so that

$$AK_n = \left[\begin{array}{c|c|c|c} Ab & A^2b & \cdots & A^nb \end{array} \right]$$

- The column space of this matrix is $A\mathcal{K}_n$. Thus the problem reduces to finding a vector $c \in \mathbb{C}^n$ such that

$$||AK_nc - b|| = minimum$$

- This can be achieved by QR factorization of AK_n. Once c is found, we can set $x_n = K_nc$.

GMRES - Generalized Minimal Residuals

- The above procedure is numerically unstable!

GMRES - Generalized Minimal Residuals

- The above procedure is numerically unstable!
- Instead use Arnoldi iteration to construct a sequence of Krylov subspace matrices, Q_n, whose columns $q_1, q_2, \ldots$ span the successive Krylov subspaces $\mathcal{K}_n$.

GMRES - Generalized Minimal Residuals

- The above procedure is numerically unstable!
- Instead use Arnoldi iteration to construct a sequence of Krylov subspace matrices, Q_n, whose columns $q_1, q_2, \ldots$ span the successive Krylov subspaces $\mathcal{K}_n$.
- Then one can write $x_n = Q_n y$ instead of $x_n = K_n c$.

GMRES - Generalized Minimal Residuals

- The above procedure is numerically unstable!
- Instead use Arnoldi iteration to construct a sequence of Krylov subspace matrices, Q_n, whose columns $q_1, q_2, \ldots$ span the successive Krylov subspaces $\mathcal{K}_n$.
- Then one can write $x_n = Q_n y$ instead of $x_n = K_n c$.
- The least squares problem reduces to finding a vector $y \in \mathbb{C}^n$ such that:

$$||AQ_n y - b|| = minimum$$

GMRES - Generalized Minimal Residuals

- Using $AQ_n = Q_{n+1}H_n$, we get:

$$||Q_{n+1}H_n y - b|| = minimum$$

GMRES - Generalized Minimal Residuals

- Using $AQ_n = Q_{n+1}H_n$, we get:

$$||Q_{n+1}H_n y - b|| = minimum$$

- Now both vectors inside the norm are in the column space of Q_{n+1}. So multiplying on the left by Q^*_{n+1} does not change that norm :

$$||H_n y - Q^*_{n+1} b|| = minimum$$

GMRES - Generalized Minimal Residuals

- Using $AQ_n = Q_{n+1}H_n$, we get:

$$||Q_{n+1}H_n y - b|| = minimum$$

- Now both vectors inside the norm are in the column space of Q_{n+1}. So multiplying on the left by Q^*_{n+1} does not change that norm :

$$||H_n y - Q^*_{n+1}b|| = minimum$$

- As b was the first vector in the column space of Q_{n+1}, $Q^*_{n+1}b = ||b||e_1$, where $e_1 = (1, 0, 0, \dots)^*$ So the problem reduces to:

$$||H_n y - ||b||e_1|| = minimum$$

GMRES - Generalized Minimal Residuals

Algorithm 1 GMRES

$q_1 = b/||b||$
for $n = 1, 2, 3, \ldots$ **do**
    $v = Aq_n$
    for $j = 1$ to n **do**
        $h_{jn} = q_j^* v$
        $v = v - h_{jn} q_j$
    end for
    $h_{n+1,n} = ||v||$
    $q_{n+1} = v/h_{n+1,n}$
    Find y to minimize $||H_n y - ||b|| e_1||$
    $x_n = Q_n y$
end for

Conjugate gradients

- Conjugate gradient iteration is the "original" Krylov subspace method – and a mainstay of scientific computing!

Conjugate gradients

- Conjugate gradient iteration is the "original" Krylov subspace method – and a mainstay of scientific computing!
- Conjugate gradient is to GMRES what Lanczos is to Arnoldi!

Conjugate gradients

- Conjugate gradient iteration is the "original" Krylov subspace method – and a mainstay of scientific computing!
- Conjugate gradient is to GMRES what Lanczos is to Arnoldi!
- It is applicable to solving $Ax = b$ problems where A is a real, positive definite, symmetric matrix.

Conjugate gradients

- Conjugate gradient iteration is the "original" Krylov subspace method – and a mainstay of scientific computing!
- Conjugate gradient is to GMRES what Lanczos is to Arnoldi!
- It is applicable to solving $Ax = b$ problems where A is a real, positive definite, symmetric matrix.
- In this case, the GMRES iteration reduces substantially and no upper Hessenberg matrix has to be constructed.

- Conjugate gradient iteration is the "original" Krylov subspace method – and a mainstay of scientific computing!
- Conjugate gradient is to GMRES what Lanczos is to Arnoldi!
- It is applicable to solving $Ax = b$ problems where A is a real, positive definite, symmetric matrix.
- In this case, the GMRES iteration reduces substantially and no upper Hessenberg matrix has to be constructed.
- Conjugate gradient can be described as: It is a system of recurrence formulas that generates the unique sequence of iterates $\{x_n \in \mathcal{K}_n\}$ with the property that at each step n $||e_n||_A$ is minimized.

Conjugate gradients

Algorithm 2 Conjugate gradients

$x_0 = 0$, $r_0 = 0$, $p_0 = r_0$
for $n = 1, 2, 3, \ldots$ **do**
  $\alpha_n = (r_{n-1}^T r_{n-1} / (p_{n-1}^T A p_{n-1})$ {Step length}
  $x_n = x_{n-1} + \alpha_n p_{n-1}$ {Approximate solution}
  $r_n = r_{n-1} - \alpha_n A p_{n-1}$ {Residual}
  $\beta_n = (r_n^T r_n)/(r_{n-1}^T r_{n-1})$ {Improvement of the step}
  $p_n = r_n + \beta_n p_{n-1}$ {Search direction}
end for

Conjugate gradients

- It can be shown that

$$r_n^T r_j = 0 \quad (j < n)$$

Conjugate gradients

- It can be shown that

$$r_n^T r_j = 0 \quad (j < n)$$

- The search directions:

$$p_n^T A p_j = 0 \quad (j < n)$$

Conjugate gradients

- It can be shown that

$$r_n^T r_j = 0 \quad (j < n)$$

- The search directions:

$$p_n^T A p_j = 0 \quad (j < n)$$

- All the x, p and r belong to the Krylov subspace:

$$\begin{aligned}\mathcal{K}_n =& \langle x_1, x_2, \ldots, x_n \rangle = \langle p_0, p_1, \ldots, p_{n-1} \rangle \\ & \langle r_0, r_1, \ldots, r_{n-1} \rangle = \langle b, Ab, \ldots A^{n-1} b \rangle\end{aligned}$$

Conjugate gradients

- Let the CG iteration be applied to a symmetric positive definite matrix problem $Ax = b$. If the iteration has not already converged, then x_n is the unique point in $\mathcal{K}_n$ that minimizes $||e_n||_A$. The convergence is monotonic:

$$||e_n||_A \leq ||e_{n-1}||_A$$

and $e_n = 0$ is achieved for some $n \leq m$.

Conjugate gradients

- We know that x_n belongs to $\mathcal{K}_n$. To show that it is unique point that minimizes $||e||_A$, consider the arbitrary point, $x = x_n - \Delta x \in \mathcal{K}_n$. The error $e = x_* - x = e_n + \Delta x$. Calculate

$$\begin{aligned} ||e||_A^2 =& (e_n + A)^T A (e_n + A) \\ =& e_n^T A e_n + (\Delta x)^T A (\Delta x) + 2 e_n^T A (\Delta x) \end{aligned}$$

Conjugate gradients

- We know that x_n belongs to $\mathcal{K}_n$. To show that it is unique point that minimizes $||e||_A$, consider the arbitrary point, $x = x_n - \Delta x \in \mathcal{K}_n$. The error $e = x_* - x = e_n + \Delta x$. Calculate

$$
\begin{aligned}
||e||_A^2 =&(e_n + A)^T A(e_n + A) \\
=&e_n^T A e_n + (\Delta x)^T A(\Delta x) + 2e_n^T A(\Delta x)
\end{aligned}
$$

- We know that the final term: $2r_n^T(\Delta x)$ is the inner product of r_n with a vector in $\mathcal{K}_n$ which has to be 0.

Conjugate gradients

- We know that x_n belongs to $\mathcal{K}_n$. To show that it is unique point that minimizes $||e||_A$, consider the arbitrary point, $x = x_n - \Delta x \in \mathcal{K}_n$. The error $e = x_* - x = e_n + \Delta x$. Calculate

$$\begin{aligned}||e||_A^2 =& (e_n + A)^T A (e_n + A) \\ =& e_n^T A e_n + (\Delta x)^T A (\Delta x) + 2 e_n^T A (\Delta x)\end{aligned}$$

- We know that the final term: $2 r_n^T (\Delta x)$ is the inner product of r_n with a vector in $\mathcal{K}_n$ which has to be 0.
- So,

$$||e||_A^2 = e_n^T A e_n + (\Delta x)^T A (\Delta x)$$

and since A is positive definite, this is minimum when $\Delta x = 0$.

Conjugate gradients

- We know that x_n belongs to $\mathcal{K}_n$. To show that it is unique point that minimizes $||e||_A$, consider the arbitrary point, $x = x_n - \Delta x \in \mathcal{K}_n$. The error $e = x_* - x = e_n + \Delta x$. Calculate

$$\begin{aligned} ||e||_A^2 =& (e_n + A)^T A (e_n + A) \\ =& e_n^T A e_n + (\Delta x)^T A (\Delta x) + 2 e_n^T A (\Delta x) \end{aligned}$$

- We know that the final term: $2 r_n^T (\Delta x)$ is the inner product of r_n with a vector in $\mathcal{K}_n$ which has to be 0.
- So,

$$||e||_A^2 = e_n^T A e_n + (\Delta x)^T A (\Delta x)$$

  and since A is positive definite, this is minimum when $\Delta x = 0$.
- The monotonicity follows from the fact that $\mathcal{K}_n \subseteq \mathcal{K}_{n+1}$.