Simple iteration procedure

Solve $Ax = b$

Known approximate solution

Preconditionning:
$$\begin{cases} M^{-1}Ax = M^{-1}b \\ M = D \longrightarrow \text{Jacobi} \\ M = L \longrightarrow \text{Gauss-Seidel} \end{cases}$$

Lower triangle

$r_0 = b - Ax_0 \longrightarrow$ residue

$M z_0 = r_0 \longrightarrow$ use of pre-conditionner

$x_k = x_{k-1} + z_{k-1} \longrightarrow$ correction

$r_k = b - Ax_k \longrightarrow$ residue

$M z_k = r_k \longrightarrow$ use of pre-conditionner

Convergence

$$e_k = A^{-1}\, b - x_k$$

$$e_k = (I - M^{-1}\, A)\, e_{k-1} = \ldots = (I - M^{-1}\, A)^k\, e_0$$

$$\lim_{k\to\infty} ||(I - M^{-1}\, A)^k|| = 0$$

$$\longleftrightarrow \quad \rho(I - M^{-1}\, A) < 1$$

Spectral radius (magnitude of largest eigenvalue)

$$Z^k\, v = \lambda^k\, v \longrightarrow 0$$

Convergence rate (from now on, we replace *$M^{-1}A$* by *A* and *$M^{-1}b$ by y*):

$$||r_{k+1}|| \leq \sqrt{1 - d^2/||A||^2}\;||r_k|| \qquad (\;0 \notin \mathcal{F}(A^H)\;)$$

$$\left|\frac{y^H\, A^H\, y}{y^H\, y}\right| \geq d$$

Field of values:

$$\mathcal{F}(A) = \{y^H A\, y \;:\; y \in \mathbf{C}^n,\; y^H y = 1\}$$

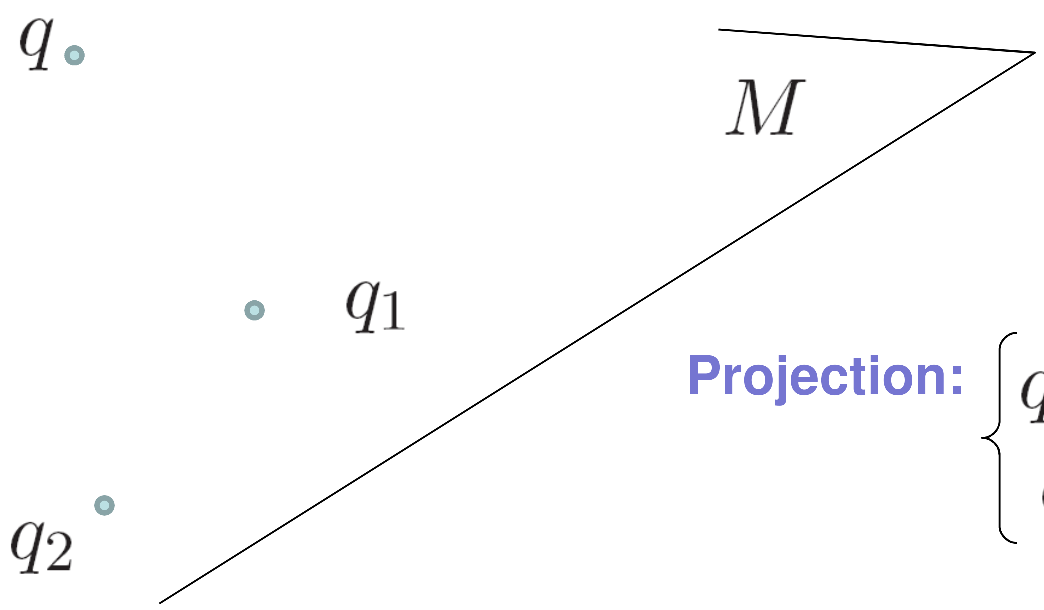
Key elements of iterative solvers

- Good pre-conditionner (often based on nearest interactions)
- Fast matrix-vector multiplication (e.g. based on Fast Multipoles)
(get far below N^2 complexity, e.g. $N \log_2 N$)
- Iterative technique with good convergence/stability properties
(get far below N iterations)

⇒ Get far below $O(N^3)$ complexity ($D\ N \log_2 N$)

Number of iterations

Minimization by projection



Projection: $\begin{cases} q_1 \in M \\ q - q_1 \perp M \end{cases}$

Distance:

$$|q - q_2|^2 = |q_2 - q_1|^2 + |q_1 - q|^2$$

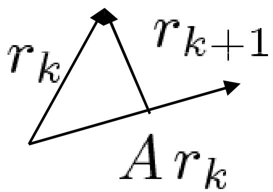
Orthomin(1) and orthomin(2)

$$x_k \in \operatorname{span}\left\{b, A\,b, \ldots, A^{k-1}\,b\right\} \longrightarrow$$ Krylov subspace

$$x_{k+1} = x_k + a_k\,(b - A\,x_k)$$

$p_k = r_k$

$$r_{k+1} = r_k - a_k\,A\,p_k \longrightarrow$$ Residue orthogonal to subspace $/\!/ A\,r_k$



$$a_k = \frac{\langle r_k, A\,r_k\rangle}{\langle A\,r_k, A\,r_k\rangle} \longrightarrow$$ Projection on $A\,r_k$

Orthomin(2):

$$\tilde{p}_k = r_k - \frac{\langle A\,r_k, A\,p_{k-1}\rangle}{\langle A\,p_{k-1}, A\,p_{k-1}\rangle}\,p_{k-1} \longrightarrow$$ Project on larger subspace

$$0 \notin \mathcal{F}(A^H)$$

Beyond that, project on all previous values of $A\,r_i$ $\left\{\begin{array}{l}\text{Orthodir(n)}\\ \text{GMRES}\end{array}\right.$

Modified Gram-Schmidt (MGS) algorithm

$$r_{11} = ||v_1||$$

$$q_1 = v_1 / r_{11}$$

$$k = 1, ..., n$$

$$\tilde{q}_k = v_k$$

$$i = 1, ..., k-1$$

$$\tilde{q}_k \longleftarrow \tilde{q}_k - \underbrace{\langle \tilde{q}_k, q_i \rangle}_{r_{ik}} \; q_i$$

$$r_{kk} = ||\tilde{q}_k||$$

$$q_k = \tilde{q}_k / r_{kk}$$

MGS defines a QR decompositon

$$v_k - \sum_{i=1}^{k-1} r_{ik}\, q_i = r_{kk}\, q_k$$

$$A = Q\,R \quad \Longrightarrow \quad \begin{cases} Ax = b \\ Q\,R\,x = b \\ R\,x = Q^H\,b \end{cases}$$

For overdetermined problems,

QR solution = least-squares solution

Normal equations: $$A^H\,A\,x = A^H\,b$$

$$\Downarrow$$

$$R\,x = Q^H\,b$$

Arnoldi algorithm: MGS on Krylov subspace

$$||q_1|| = 1 \qquad\qquad \text{span}\left\{r_0, A\,r_0, \ldots, A^k\,r_0\right\}$$

normalize

$$j = 1, 2, \ldots,$$

$$\tilde{q}_{j+1} = A\,q_j$$

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

$$h_{ij} = \langle \tilde{q}_{j+1}, q_i \rangle$$

$$\tilde{q}_{j+1} \longleftarrow \tilde{q}_{j+1} - h_{ij}\,q_i$$

$$h_{j+1,j} = ||\tilde{q}_{j+1}||$$

$$q_{j+1} = \tilde{q}_{j+1} / h_{j+1,j}$$

Simple iteration procedure

Solve $Ax = b$

Preconditionning: $\left\{ \begin{array}{l} M^{-1} A x = M^{-1} b \\ M = D \longrightarrow \text{Jacobi} \\ M = L \longrightarrow \text{Gauss-Seidel} \end{array} \right.$

Known approximate solution

Lower triangle

$$r_0 = b - A x_0 \longrightarrow \text{residue}$$

$$M z_0 = r_0 \longrightarrow \text{use of pre-conditionner}$$

$$x_k = x_{k-1} + z_{k-1} \longrightarrow \text{correction}$$

$$r_k = b - A x_k \longrightarrow \text{residue}$$

$$M z_k = r_k \longrightarrow \text{use of pre-conditionner}$$

Formation of Krylov subspace

$$x'_k = x_k - x_o$$

If A already preconditionned:

$$x'_k = x'_{k-1} + (b - A\,x_{k-1})$$

$$x'_k = x'_{k-1} + (b - A\,x_o - A\,x'_{k1})$$

$$x'_k = r_o + (I - A\,x'_{k-1})$$

$$x'_1 = r_o$$

$$x'_k \in \ \mathrm{span}\left\{r_0, A\,r_0, \ldots, A^k\,r_0\right\}$$

Primitive GMRES (Generalized Minimum Residual)

Principle: the solution belongs to the (orthogonalized) Krylov subspace…

$$x_k = x_\circ + Q_k\, y_k$$

…with coefficients y_k obtained in the least-square sense

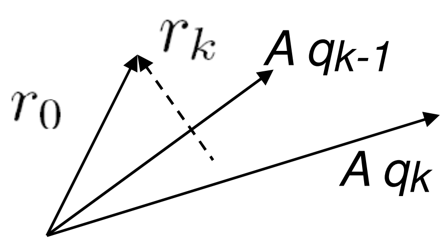
…for instance with the help of a QR solution !

The residue: $b - A\, x_k = b - A\, x_\circ - A\, Q_k\, y_k$

$$r_k = r_0 - A\, Q_k\, y_k$$

is minimized if orthogonal to the subspace

formed by $A\, Q_k$



Primitive FOM (Full orthogonalization method)

Principle: the solution belongs to the (orthogonalized) Krylov subspace…

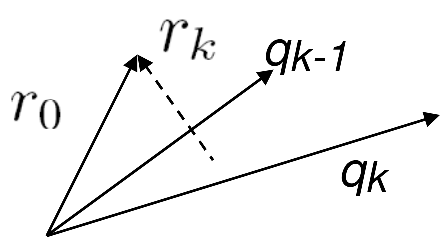
$$x_k = x_\circ + Q_k\, y_k$$

…with coefficients y_k obtained such that (**for SDP A matrices**) the A-norm of the error gets minimized:

$$\epsilon = (\tilde{x} - x)^H A\, (\tilde{x} - x)$$

$$\tilde{x} \longrightarrow \tilde{x} + \delta_v \qquad \delta_v \in \operatorname{span}\left\{r_0, A\, r_0, ..., A^k\, r_0\right\}$$

$$\delta_v^H\ \left(A\,(\tilde{x} - x)\right) + \delta_v^H\ A\ \delta_v > 0$$



« Brutal » algorithm for FOM and GMRES

```
q1=ro;q=q1; % hyp: xo=0, q = Krylov
stop=0; k=0;

while (stop==0)
k=k+1;

p1=A*q1; % matrix-vector product
q=[q,p1]; %  A*Kyrlov subspace

qa=q(:,1:k);   % Kylov subspace K
qb=q(:,2:k+1);   % A*Krylov subspace
[qa,rr]=qr(qa,0);   % Orthogonalize K
qb=(rr'\qb')';  % same transformation to A*K
```

We should not orthogonalize at each step !

```
if (gmres)
   [qq,rr]=qr(qb,0); % GMRES
   % has been simplified w.r.t. course
   y=rr\(qq'*ro);
elseif (fom)
   [qq,rr]=qr(qa'*qb,0); % FOM
   y=rr\(qq'*qa'*ro);
end

x=qa*y;
res=norm(qb*y-ro)/norm(ro) % residue
if ((res<1e-12)||(k>20)) stop=1;
end

end
```

Arnoldi algorithm: Upper Hessenberg Matrix

$$||q_1|| = 1$$

$$j = 1, 2, ...,$$

$$\tilde{q}_{j+1} = A\, q_j$$

$$i = 1, ..., j$$

$$h_{ij} = \langle \tilde{q}_{j+1}, q_i \rangle$$

$$\tilde{q}_{j+1} \longleftarrow \tilde{q}_{j+1} - h_{ij}\, q_i$$

$$h_{j+1,j} = ||\tilde{q}_{j+1}||$$

$$q_{j+1} = \tilde{q}_{j+1} / h_{j+1,j}$$

$$\boldsymbol{H_{43}} = \begin{array}{c} \boldsymbol{H_{33}} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ 0 & h_{32} & h_{33} \end{bmatrix} \\ \begin{matrix} 0 & \quad 0 & \quad h_{43} \end{matrix} \end{array}$$

Arnoldi recursion

From orthogonalization algorithm:

$$A\,q_j - \sum_{i=1}^{j} h_{ij}\,q_i = h_{j+1,j}\;q_{j+1}$$

Recursion formula:

$$\underset{\text{nxn}}{A}\,\underset{\text{nxk}}{Q_k} = Q_k\,H_k + \underbrace{\left[0_{n\times k\text{-}1} h_{k+1,k}\,q_{k+1}\right]}_{h_{k+1,k}\,q_{k+1}\,e_k^T} = \underset{\text{nx(k+1)}}{Q_{k+1}}\,\underset{\text{(k+1)xk}}{H_{k+1,k}}$$

Arnoldi recursion applied to FOM

Preliminary: $$Q_m^T\, A\, Q_m = H_m$$

Orthogonality w.r.t. Q: $$Q_m^T\, r_m = Q_m^T\,\left(b - A\,(x_o + Q_m\, y_m)\right) = 0$$

$$Q_m^T\, r_o = Q_m^T\, A\, Q_m\, y_m$$

$$\beta\, e_1 = H_m\, y_m$$

$||r_0||$

$$y_m = H_m^{-1}\,(\beta\, e_1)$$

Residue:

$$\beta\, q_1 - A\, Q_m\, y_m = \cancel{\beta\, q_1} - \cancel{Q_m\, H_m\, y_m} - h_{m+1,m}\, q_{m+1}\, e_m^T\, y_m$$

GMRES: residue minimization

$$\min_y ||r_0 - A\,Q_k\,y|| = \min_y ||r_0 - Q_{k+1}\,H_{k+1,k}\,y|| =$$

$$\min_y ||Q_{k+1}\,(\beta\,e_1 - H_{k+1,k}\,y)|| = \min_y ||\beta\,e_1 - H_{k+1,k}\,y||$$

$\beta = ||r_0||$

All 0's, except entry 1

as proven earlier

$$||b - A\,x_k|| = ||\beta\,e_1 - \underbrace{F^H\,R}\,y^k|| = ||\beta F e_1 - R\,y_k||$$

QR decomposition of $H_{k+1,k}$

Succession of orthonormal transformations:

$FF^H = I$

k+1 comp. of $\beta||F\zeta_1||$

Recursive orthogonal transformations for H

Orthogonalisation realised
at previous iteration:

$$(F_k \ldots F_1)\ H_{k+1,k} = \begin{bmatrix} x & x & \ldots & x \\ 0 & x & \ldots & x \\ \ldots & \ldots & \ldots & x \\ 0 & \ldots & 0 & x \\ 0 & \ldots & \ldots & 0 \end{bmatrix}$$

Same operations on new
Hessenberg matrix:

$$(F_k \ldots F_1)\ H_{k+2,k+1} = \begin{bmatrix} x & x & \ldots & x & x \\ 0 & x & \ldots & x & x \\ \ldots & \ldots & \ldots & x & x \\ 0 & \ldots & 0 & x & x \\ 0 & \ldots & \ldots & 0 & d \\ 0 & \ldots & \ldots & 0 & h \end{bmatrix}$$

Find new
operation to
null that entry

from same
operations applied
to new column of H

Givens rotations

$$F_i = \begin{bmatrix} I & & & \\ & c_i & s_i & \\ & -s_i^\star & c_i & \\ & & & I \end{bmatrix}$$

$$c = |d| / \sqrt{|d|^2 + |h|^2} \quad s^\star = c\, h / d$$ if d≠0

$$c = 0 \quad s = 1$$ if d=0

Expression of residual

$$||b - A\, x_k||^2 = ||\beta\, \zeta_1 - F^H\, R\, y^k||^2 = ||\beta F \zeta_1 - R\, y_k||^2$$

$=$ Norm^2 of k+1 comp. of $F \zeta_1$

Conjugate Gradients for symmetric DP matrices

$$H_m = Q_m^T\, A\, Q_m$$ **must be symmetric** $$H_m = T_m$$

Tri-diagonal !!!

$$j = 1, 2, ...,$$

$$\beta_j = h_{j-1,j} = \langle q_{j-1}, A\, q_j \rangle$$

$$\tilde{q}_j \longleftarrow A\, q_j - \beta_j\, q_{j-1}$$

$$\alpha_j = \langle \tilde{q}_j, q_j \rangle$$

$$\tilde{q}_j \longleftarrow \tilde{q}_j - \alpha_j\, q_j$$

$$\beta_{j+1} = ||\tilde{q}_j||$$

$$q_{j+1} = \tilde{q}_j / \beta_{j+1}$$

General Arnoldi:

$$j = 1, 2, ...,$$

$$\tilde{q}_{j+1} = A\, q_j$$

$$i = 1, ..., j$$

$$h_{ij} = \langle \tilde{q}_{j+1}, q_i \rangle$$

$$\tilde{q}_{j+1} \longleftarrow \tilde{q}_{j+1} - h_{ij}\, q_i$$

$$h_{j+1,j} = ||\tilde{q}_{j+1}||$$

$$q_{j+1} = \tilde{q}_{j+1} / h_{j+1,j}$$

Conjugate Gradients for symmetric DP matrices

$H_m = Q_m^T \, A \, Q_m$ **must be symmetric** $H_m = T_m$

Tri-diagonal

Solution: $y_m = T_m^{-1} (\beta \, e_1)$

Arnoldi's 3-terms recurrence:

$$\beta_{j+1} \, q_{j+1} = A \, q_j - \alpha_j \, q_j - \beta_{j-1} \, q_{j-1}$$