Scientific Computing TU Berlin Winter 2021/22 © Jürgen Fuhrmann
Notebook 17

```
begin
    using PlutoUI
    using HypertextLiteral
end
```

Generalization of iteration schemes

Generalization of iteration schemes

- So far we considered simple iterative schemes, perhaps with preconditioners
- Here, we introduce methods which indeed in many cases yield faster convergence than simple iterative schemes
- Material after
  - M. Gutknecht **A Brief Introduction to Krylov Space Methods for Solving Linear Systems**
  - J. Shewchuk **Introduction to the Conjugate Gradient Method Without the Agonizing Pain**
- Extended reading: J.Liesen, Z. Strakoš: Krylov Subspace Methods: Principles and Analysis
- Extended coverage of the topic available at TU: Prof. Jörg Liesen, Prof. Reinhard Nabben are active researchers in the field.

A reinterpretation of the simple iteration method

Solve $Au = b$ with tolerance ε:

1. Choose initial value u_0, set $k = 0$
2. Calculate *residuum* $r_k = Au_k - b$
3. Test convergence: if $||r_k|| < \varepsilon$ set $u = u_k$, finish
4. Calculate *update*: $v_k = \alpha r_k$
5. Update solution: $u_{k+1} = u_k - v_k$, set $k = k + 1$, repeat with step 2.

- From step 4:

$$\begin{aligned} Au_{k+1} &= Au_k - \alpha Ar_k \\ Au_{k+1} - b &= Au_k - b - \alpha Ar_k \\ r_{k+1} &= r_k - \alpha Ar_k \end{aligned}$$

- Therefore for the residual, we obtain $r_k = p_k(A)r_0 \in \text{span}\{r_0, Ar_0, \ldots, A^k r_0\}$ where $p(\xi) = (1-\alpha\xi)^k$ is a polynomial of degre k
- For the iterate u_k, we have

$$\begin{aligned} u_k &= u_{k-1} - \alpha r_{k-1} = u_{k-2} - \alpha r_{k-2} - \alpha r_{k-1} \\ &= u_0 - \alpha(r_0 + r_1 + \cdots + r_{k-1}) \\ &= u_0 + q_{k-1}(A)r_0 \end{aligned}$$

where q_{k-1} is a polynomial of degree $k-1$.

- From $r_k = Au_k - b = Au_0 - b + Aq_{k-1}(A)r_0 = (I + Aq_{k-1}(A))r_0$ one obtains $p_k(\xi) = 1 + \xi q_k(\xi)$.
- Consequently, $u_k \in u_0 + \text{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$

Definition: Let $A \in \mathbb{R}^{N\times N}$ be nonsingular, let $0 \neq y \in \mathbb{R}^n$. The k-th *Krylov subspace* generated from A by y is defined as $\mathcal{K}_k(A, y) = \text{span}\{y, Ay, \ldots, A^{k-1}y\}$.

Obviously,

- $\mathcal{K}_1 \subseteq \mathcal{K}_2 \subseteq \cdots \subseteq \mathcal{K}_k$
- $\dim \mathcal{K}_i \leq \dim \mathcal{K}_{i-1} + 1$

For the simple iteration, we established

$$\begin{aligned} u_k &= u_0 + q_{k-1}(A)r_0 \in \mathcal{K}_k(A, r_0) \\ r_k &= p_k(A)r_0 \in \mathcal{K}_{k+1}(A, r_0) \\ p_k(\xi) &= 1 + \xi q_k(\xi) \\ p_k(0) &= 1 \end{aligned}$$

with particular polynomials p, q.

Are these the best possible ones ?

Definition: Let $A \in \mathbb{R}^{N\times N}$ be nonsingular, let $0 \neq y \in \mathbb{R}^N$. An iterative method such that

$$u_k = u_0 + q_{k-1}(A)r_0 \in \mathcal{K}_k(A, r_0)$$

where q_{k-1} is a polynomial of degree k is called *Krylov subspace method*.

The case of symmetric positive definite matrices

Assume A is symmetric, positive definite (spd). Define the quadratic form

$$a(u,v) = (Au, v) = v^T Au = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} v_i u_j$$

As A is spd, for all $u \neq 0$ we have $(Au, u) > 0$.

For a given vector b, regard the function

$$f(u) = \frac{1}{2} a(u,u) - b^T u$$

What is the minimizer of f ?

$$f'(u) = Au - b = 0$$

$\Rightarrow$ Solution of SPD system $\equiv$ minimization of f.

The method of steepest descent

- Given some vector u_i, look for a new iterate u_{i+1}.
- The direction of steepest descend is given by $-f'(u_i)$.
- So look for u_{i+1} in the direction of $-f'(u_i) = r_i = b - Au_i$ such that it minimizes f in this direction, i.e. set $u_{i+1} = u_i + \alpha r_i$ with α choosen from

$$\begin{aligned} 0 &= \frac{d}{d\alpha} f(u_i + \alpha r_i) = f'(u_i + \alpha r_i) \cdot r_i \\ &= (b - A(u_i + \alpha r_i), r_i) \\ &= (b - Au_i, r_i) - \alpha (Ar_i, r_i) \\ &= (r_i, r_i) - \alpha (Ar_i, r_i). \end{aligned}$$

Consequently,

$$\alpha = \frac{(r_i, r_i)}{(Ar_i, r_i)}.$$

The resulting iteration scheme is

$$\begin{aligned} r_i &= b - Au_i \\ \alpha_i &= \frac{(r_i, r_i)}{(Ar_i, r_i)} \\ u_{i+1} &= u_i + \alpha_i r_i \end{aligned}$$

and can be shown to belong to the class of Krylov subspace methods.

Let $\hat{u}$ the exact solution. Define $e_i = u_i - \hat{u}$, then $r_i = -Ae_i$

Let $||u||_A = (Au, u)^{\frac{1}{2}}$ be the *energy norm* wrt. A.

Theorem The convergence rate of the steepest descent method can be estimated as

$$||e_i||_A \leq \left(\frac{\kappa - 1}{\kappa + 1}\right)^i ||e_0||_A$$

where $\kappa = \frac{\lambda_{max}(A)}{\lambda_{min}(A)}$ is the spectral condition number.

The simple Richardson iteration $u_{k+1} = u_k - \alpha(Au_k - f)$ needs good eigenvalue estimate to be optimal with $\alpha = \frac{2}{\lambda_{max}+\lambda_{min}}$

- In this optimal case, the asymptotic convergence rate is $\rho = \frac{\kappa-1}{\kappa+1}$
- Steepest descent has the same rate without need for spectral estimate

Can we get a better estimate ? E.g. with $\sqrt{\kappa}$ instead of κ as discussed in the previous lecture ?

The method of conjugate directions

For steepest descent, there is no guarantee that a search direction $d_i = r_i = -Ae_i$ is not used several times. If all search directions would be orthogonal, or, indeed, A-orthogonal, one could control this situation.

So, let $d_0, d_1 \dots d_{n-1}$ be a series of A-orthogonal (or conjugate) search directions, i.e. $(Ad_i, d_j) = 0$, $i \neq j$.

- Look for u_{i+1} in the direction of d_i such that it minimizes f in this direction, i.e. set $u_{i+1} = u_i + \alpha_i d_i$ with α choosen from

$$\begin{aligned} 0 &= \frac{d}{d\alpha} f(u_i + \alpha d_i) = f'(u_i + \alpha d_i) \cdot d_i \\ &= (b - A(u_i + \alpha d_i), d_i) \\ &= (b - Au_i, d_i) - \alpha(Ad_i, d_i) \\ &= (r_i, d_i) - \alpha(Ad_i, d_i) \\ \alpha_i &= \frac{(r_i, d_i)}{(Ad_i, d_i)} \end{aligned}$$

$\Rightarrow u_{i+1} \in \text{span}\{d_0 \dots d_i\}$

Conjugate gradients

- Choose $d_0 \dots d_i$ such that $\text{span}\{d_0 \dots d_i\} = \mathcal{K}_i(A, r_0)$.
- Orthogonalize by Gram-Schmidt
- It can be shown (and probably has been shown in another course) that
  - in this particular case of spd matrices the process results in short recursions, i.e. d_{i+1} depends only on d_i and not on all $d_0 \dots d_i$
  - $u_i \in u_0 + \mathcal{K}_i(A, r_0)$ minimizes the energy norm $||e_i||_A = (Ae_i, e_i)$ of the error with respect to the elements of the Krylov subspace
  - $r_{i+1} \perp \mathcal{K}_i(A, r_0)$
  - There are at most N directions, so the method yields the exact solution after at most N iteration steps.

Given initial value u_0, spd matrix A, right hand side b, the algorithm reads as

$$\begin{aligned} d_0 &= r_0 = b - Au_0 \\ \alpha_i &= \frac{(r_i, r_i)}{(Ad_i, d_i)} \\ u_{i+1} &= u_i + \alpha_i d_i \\ r_{i+1} &= r_i - \alpha_i Ad_i \\ \beta_{i+1} &= \frac{(r_{i+1}, r_{i+1})}{(r_i, r_i)} \\ d_{i+1} &= r_{i+1} + \beta_{i+1} d_i \end{aligned}$$

At the i-th step, the algorithm yields the element from $e_0 + \mathcal{K}_i$ with the minimum energy error.

Theorem The convergence rate of the method is

$$||e_i||_A \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^i ||e_0||_A$$

where $\kappa = \frac{\lambda_{max}(A)}{\lambda_{min}(A)}$ is the spectral condition number of A.

Let M be spd, and spectrally equivalent to A, and assume that $\kappa(M^{-1}A) << \kappa(A)$.

Let E be such that $M = EE^T$. Then, $\sigma(M^{-1}A) = \sigma(E^{-1}AE^{-T})$:

Assume $M^{-1}Au = \lambda u$. We have

$$(E^{-1}AE^{-T})(E^Tu) = (E^TE^{-T})E^{-1}Au = E^TM^{-1}Au = \lambda E^Tu$$

$\Leftrightarrow E^Tu$ is an eigenvector of $E^{-1}AE^{-T}$ with eigenvalue λ.

Now we can use the CG algorithm for the preconditioned system

$$E^{-1}AE^{-T}\tilde{x} = E^{-1}b$$

with $\tilde{u} = E^Tu$

$$\begin{aligned}
\tilde{d}_0 &= \tilde{r}_0 = E^{-1}b - E^{-1}AE^{-T}u_0 \\
\alpha_i &= \frac{(\tilde{r}_i, \tilde{r}_i)}{(E^{-1}AE^{-T}\tilde{d}_i, \tilde{d}_i)} \\
\tilde{u}_{i+1} &= \tilde{u}_i + \alpha_i\tilde{d}_i \\
\tilde{r}_{i+1} &= \tilde{r}_i - \alpha_iE^{-1}AE^{-T}\tilde{d}_i \\
\beta_{i+1} &= \frac{(\tilde{r}_{i+1}, \tilde{r}_{i+1})}{(\tilde{r}_i, \tilde{r}_i)} \\
\tilde{d}_{i+1} &= \tilde{r}_{i+1} + \beta_{i+1}\tilde{d}_i
\end{aligned}$$

Not very practical as we need E

Assume $\tilde{r}_i = E^{-1}r_i, \tilde{d}_i = E^Td_i$, we get the equivalent algorithm

$$\begin{aligned}
r_0 &= b - Au_0 \\
d_0 &= M^{-1}r_0 \\
\alpha_i &= \frac{(M^{-1}r_i, r_i)}{(Ad_i, d_i)} \\
u_{i+1} &= u_i + \alpha_id_i \\
r_{i+1} &= r_i - \alpha_iAd_i \\
\beta_{i+1} &= \frac{(M^{-1}r_{i+1}, r_{i+1})}{(r_i, r_i)} \\
d_{i+1} &= M^{-1}r_{i+1} + \beta_{i+1}d_i
\end{aligned}$$

It relies on the solution of the preconditioning system, the calculation of the matrix vector product and the calculation of the scalar product.

Theorem: The convergence rate of the method can be estimated from

$$||e_i||_{E^{-1}AE^{-T}} \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^i ||e_0||_{E^{-1}AE^{-T}}$$

where $\kappa = \frac{\gamma_{max}}{\gamma_{min}}$ comes from $\gamma_{min}(Mu, u) \leq (Au, u) \leq \gamma_{max}(Mu, u)$.

- Usually we stop the iteration when the residual r becomes small. However during the iteration, floating point errors occur which distort the calculations and lead to the fact that the accumulated residuals

$$r_{i+1} = r_i - \alpha_i A d_i$$

  give a much more optimistic picture on the state of the iteration than the real residual

$$r_{i+1} = b - A u_{i+1}$$

- The convergence rate estimate in terms of $\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$ indeed provides a qualitatively better complexity estimate for the solution algorithm
- Always consider CG when solving symmetric positive definite linear systems iteratively for PDE problems in dimension 2 and higher

Unsymmetric problems

Search the new iterate

$$u_k = u_0 + q_{k-1}(A) r_0 \in \mathcal{K}_k(A, r_0)$$

such that $r_k = ||Au_k - b||$ is minimized. This results in the *Generalized Minimum Residual* (GMRES) method.

- In order to find a good solution of this problem, we need to find an orthogonal basis of $\mathcal{K}_k \Rightarrow$ run an orthogonalization algorithm at each step
- One needs to store at least k vectors simultaneously $\Rightarrow$ usually, the iteration is restarted after a fixed number of iteration steps to keep the dimension of $\mathcal{K}_k$ limited
- There are preconditioned variants
- For symmetric matrices, we once again (as for CG) we get short three-term recursions, and there is no need to store a full Krylov basis. This results in the MINRES method
- Choosing p_k and q_k such that we get short recursions always will sacrifice some of the convergence estimates for GMRES. Nevertheless, this appraoch is tried quite often, resuling in particular in the BiCGstab and CGS metods.