Example slides from a few `Prosper`-based lectures

Matthias Heil & Andrew L. Hazel

.

M.Heil@maths.man.ac.uk & ahazel@maths.man.ac.uk

http://www.maths.man.ac.uk/~mheil

Department of Mathematics

University of Manchester

What are Krylov subspaces and why do they matter?

We want to solve

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

What are Krylov subspaces and why do they matter?

We want to solve

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

- Consider the elementary splitting:

$$\mathbf{I}\mathbf{x} = (\mathbf{I} - \mathbf{A})\mathbf{x} + \mathbf{b}$$

What are Krylov subspaces and why do they matter?

We want to solve

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

- Consider the elementary splitting:

$$\mathbf{I}\mathbf{x} = (\mathbf{I} - \mathbf{A})\mathbf{x} + \mathbf{b}$$

- Write this as a fixpoint iteration:

$$\mathbf{x}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{x}_n + \mathbf{b}$$

What are Krylov subspaces and why do they matter?

We want to solve

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

- Consider the elementary splitting:

$$\mathbf{I}\mathbf{x} = (\mathbf{I} - \mathbf{A})\mathbf{x} + \mathbf{b}$$

- Write this as a fixpoint iteration:

$$\mathbf{x}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{x}_n + \mathbf{b}$$

- Rewrite as

$$\mathbf{x}_{n+1} = \mathbf{x}_n + (\mathbf{b} - \mathbf{A}\mathbf{x}_n)$$

What are Krylov subspaces and why do they matter?

We want to solve

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

- Consider the elementary splitting:

$$\mathbf{I}\mathbf{x} = (\mathbf{I} - \mathbf{A})\mathbf{x} + \mathbf{b}$$

- Write this as a fixpoint iteration:

$$\mathbf{x}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{x}_n + \mathbf{b}$$

- Rewrite as

$$\mathbf{x}_{n+1} = \mathbf{x}_n + (\mathbf{b} - \mathbf{A}\mathbf{x}_n)$$

- or:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{r}_n$$

  where $\mathbf{r}_n = \mathbf{b} - \mathbf{A}\mathbf{x}_n$ is the residual.

What are Krylov subspaces and why do they matter?

We want to solve

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

- Consider the elementary splitting:

$$\mathbf{I}\mathbf{x} = (\mathbf{I} - \mathbf{A})\mathbf{x} + \mathbf{b}$$

- Write this as a fixpoint iteration:

$$\mathbf{x}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{x}_n + \mathbf{b}$$

- Rewrite as

$$\mathbf{x}_{n+1} = \mathbf{x}_n + (\mathbf{b} - \mathbf{A}\mathbf{x}_n)$$

- or:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{r}_n$$

  where $\mathbf{r}_n = \mathbf{b} - \mathbf{A}\mathbf{x}_n$ is the residual.

- For simplicity assume $\mathbf{x}_0 = \mathbf{0}$, then:

$$\mathbf{x}_{n+1} = \sum_{i=0}^{n} \mathbf{r}_i \quad \text{where } \mathbf{r}_0 = \mathbf{b}.$$

What are Krylov subspaces and why do they matter? (2)

$$\mathbf{x}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{x}_n + \mathbf{b} \qquad (1)$$

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{r}_n \qquad (2)$$

$$\mathbf{x}_{n+1} = \sum_{i=0}^{n} \mathbf{r}_i \qquad (3)$$

What are Krylov subspaces and why do they matter? (2)

$$\mathbf{x}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{x}_n + \mathbf{b} \quad (1) \qquad \mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{r}_n \quad (2) \qquad \mathbf{x}_{n+1} = \sum_{i=0}^{n} \mathbf{r}_i \quad (3)$$

- Multiply (2) by $-\mathbf{A}$ and add $\mathbf{b}$:

$$\mathbf{b} - \mathbf{A}\mathbf{x}_{n+1} = \mathbf{b} - \mathbf{A}\mathbf{x}_n - \mathbf{A}\mathbf{r}_n$$

What are Krylov subspaces and why do they matter? (2)

$$\mathbf{x}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{x}_n + \mathbf{b} \quad (1) \qquad \mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{r}_n \quad (2) \qquad \mathbf{x}_{n+1} = \sum_{i=0}^{n} \mathbf{r}_i \quad (3)$$

- Multiply (2) by $-\mathbf{A}$ and add $\mathbf{b}$:

$$\mathbf{b} - \mathbf{A}\mathbf{x}_{n+1} = \mathbf{b} - \mathbf{A}\mathbf{x}_n - \mathbf{A}\mathbf{r}_n$$

$$\mathbf{r}_{n+1} = \mathbf{r}_n - \mathbf{A}\mathbf{r}_n$$

What are Krylov subspaces and why do they matter? (2)

$$\mathbf{x}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{x}_n + \mathbf{b} \qquad (1) \qquad\qquad \mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{r}_n \qquad (2) \qquad\qquad \mathbf{x}_{n+1} = \sum_{i=0}^{n} \mathbf{r}_i \qquad (3)$$

- Multiply (2) by $-\mathbf{A}$ and add $\mathbf{b}$:

$$\mathbf{b} - \mathbf{A}\mathbf{x}_{n+1} = \mathbf{b} - \mathbf{A}\mathbf{x}_n - \mathbf{A}\mathbf{r}_n$$

$$\mathbf{r}_{n+1} = \mathbf{r}_n - \mathbf{A}\mathbf{r}_n$$

$$\mathbf{r}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{r}_n$$

What are Krylov subspaces and why do they matter? (2)

$$\mathbf{x}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{x}_n + \mathbf{b} \quad (1) \qquad \mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{r}_n \quad (2) \qquad \mathbf{x}_{n+1} = \sum_{i=0}^{n} \mathbf{r}_i \quad (3)$$

- Multiply (2) by $-\mathbf{A}$ and add $\mathbf{b}$:

$$\mathbf{b} - \mathbf{A}\mathbf{x}_{n+1} = \mathbf{b} - \mathbf{A}\mathbf{x}_n - \mathbf{A}\mathbf{r}_n$$

$$\mathbf{r}_{n+1} = \mathbf{r}_n - \mathbf{A}\mathbf{r}_n$$

$$\mathbf{r}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{r}_n$$

$$\mathbf{r}_{n+1} = (\mathbf{I} - \mathbf{A})^{n+1}\mathbf{r}_0$$

Note that this shows that the iteration only converges if $\rho(\mathbf{I} - \mathbf{A}) < 1$.

What are Krylov subspaces and why do they matter? (2)

$$\mathbf{x}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{x}_n + \mathbf{b} \qquad (1) \qquad\qquad \mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{r}_n \qquad (2) \qquad\qquad \mathbf{x}_{n+1} = \sum_{i=0}^{n} \mathbf{r}_i \qquad (3)$$

- Multiply (2) by $-\mathbf{A}$ and add $\mathbf{b}$:

$$\mathbf{b} - \mathbf{A}\mathbf{x}_{n+1} = \mathbf{b} - \mathbf{A}\mathbf{x}_n - \mathbf{A}\mathbf{r}_n$$

$$\mathbf{r}_{n+1} = \mathbf{r}_n - \mathbf{A}\mathbf{r}_n$$

$$\mathbf{r}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{r}_n$$

$$\mathbf{r}_i = (\mathbf{I} - \mathbf{A})^i \mathbf{r}_0 \qquad (4)$$

What are Krylov subspaces and why do they matter? (2)

$$\mathbf{x}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{x}_n + \mathbf{b} \qquad (1) \qquad\qquad \mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{r}_n \qquad (2) \qquad\qquad \mathbf{x}_{n+1} = \sum_{i=0}^{n} \mathbf{r}_i \qquad (3)$$

- Multiply (2) by $-\mathbf{A}$ and add $\mathbf{b}$:

$$\mathbf{b} - \mathbf{A}\mathbf{x}_{n+1} = \mathbf{b} - \mathbf{A}\mathbf{x}_n - \mathbf{A}\mathbf{r}_n$$

$$\mathbf{r}_{n+1} = \mathbf{r}_n - \mathbf{A}\mathbf{r}_n$$

$$\mathbf{r}_{n+1} = (\mathbf{I} - \mathbf{A})\mathbf{r}_n$$

$$\mathbf{r}_i = (\mathbf{I} - \mathbf{A})^i \mathbf{r}_0 \qquad (4)$$

- Insert (4) into (3):

$$\mathbf{x}_{n+1} = \sum_{i=0}^{n} (\mathbf{I} - \mathbf{A})^i \mathbf{r}_0 = \sum_{i=0}^{n} (\mathbf{I} - \mathbf{A})^i \mathbf{b}$$

What are Krylov subspaces and why do they matter? (3)

- So while we're carrying out the iteration

$$\mathbf{x}_n = (\mathbf{I} - \mathbf{A})\mathbf{x}_{n-1} + \mathbf{b}$$

the method generates the iterates

$$\mathbf{x}_n = \sum_{i=0}^{n-1} (\mathbf{I} - \mathbf{A})^i \mathbf{b}$$

- So while we're carrying out the iteration

$$\mathbf{x}_n = (\mathbf{I} - \mathbf{A})\mathbf{x}_{n-1} + \mathbf{b}$$

the method generates the iterates

$$\mathbf{x}_n = \sum_{i=0}^{n-1} (\mathbf{I} - \mathbf{A})^i \mathbf{b}$$

- Expanding this gives

$$\mathbf{x}_n = \sum_{i=0}^{n-1} \alpha_i \mathbf{A}^i \mathbf{b}$$

for some coefficients α_i.

What are Krylov subspaces and why do they matter? (3)

- So while we're carrying out the iteration

$$\mathbf{x}_n = (\mathbf{I} - \mathbf{A})\mathbf{x}_{n-1} + \mathbf{b}$$

  the method generates the iterates

$$\mathbf{x}_n = \sum_{i=0}^{n-1} (\mathbf{I} - \mathbf{A})^i \mathbf{b}$$

- Expanding this gives

$$\mathbf{x}_n = \sum_{i=0}^{n-1} \alpha_i \mathbf{A}^i \mathbf{b}$$

  for some coefficients α_i.

- In other words:

$$\mathbf{x}_n \in span(\mathbf{b}, \mathbf{A}\mathbf{b}, \mathbf{A}^2\mathbf{b}, ..., \mathbf{A}^{n-1}\mathbf{b}) =: \mathcal{K}_n(\mathbf{A}, \mathbf{b})$$

  where $\mathcal{K}_n(\mathbf{A}, \mathbf{b})$ is the Krylov subspace.

What are Krylov subspaces and why do they matter? (4)

- Our (well, Richardson's...) method produces:

$$\mathbf{x}_n = \sum_{i=0}^{n-1} \alpha_i \mathbf{A}^i \mathbf{b} \quad \in \quad \mathcal{K}_n(\mathbf{A}, \mathbf{b})$$

for some coefficients α_i.

What are Krylov subspaces and why do they matter? (4)

- Our (well, Richardson's...) method produces:

$$\mathbf{x}_n = \sum_{i=0}^{n-1} \alpha_i \mathbf{A}^i \mathbf{b} \quad \in \quad \mathcal{K}_n(\mathbf{A}, \mathbf{b})$$

  for some coefficients α_i.

- Are there any 'better' choices for the coefficients α_i?

What are Krylov subspaces and why do they matter? (4)

- Our (well, Richardson's...) method produces:

$$\mathbf{x}_n = \sum_{i=0}^{n-1} \alpha_i \mathbf{A}^i \mathbf{b} \quad \in \quad \mathcal{K}_n(\mathbf{A}, \mathbf{b})$$

  for some coefficients α_i.

- Are there any 'better' choices for the coefficients α_i?
- Yes, but it depends how you define 'better', or indeed 'best' as we're trying to find the 'best' approximation in a given Krylov subspace.

What are Krylov subspaces and why do they matter? (4)

- Our (well, Richardson's...) method produces:

$$\mathbf{x}_n = \sum_{i=0}^{n-1} \alpha_i \mathbf{A}^i \mathbf{b} \quad \in \quad \mathcal{K}_n(\mathbf{A}, \mathbf{b})$$

  for some coefficients α_i.

- Are there any 'better' choices for the coefficients α_i?
- Yes, but it depends how you define 'better', or indeed 'best' as we're trying to find the 'best' approximation in a given Krylov subspace.
- Various approaches and the resulting schemes:
  - Minimise the residual (MINRES, GMRES)
  - Minimise the error (GMERR)
  - Make the residual orthogonal to the Krylov subspace (Ritz/Galerkin: CG, CGS)
  - ...

What are Krylov subspaces and why do they matter? (4)

- Our (well, Richardson's...) method produces:

$$\mathbf{x}_n = \sum_{i=0}^{n-1} \alpha_i \mathbf{A}^i \mathbf{b} \quad \in \quad \mathcal{K}_n(\mathbf{A}, \mathbf{b})$$

  for some coefficients α_i.

- Are there any 'better' choices for the coefficients α_i?
- Yes, but it depends how you define 'better', or indeed 'best' as we're trying to find the 'best' approximation in a given Krylov subspace.
- Various approaches and the resulting schemes:
  - Minimise the residual (MINRES, GMRES)
  - Minimise the error (GMERR)
  - Make the residual orthogonal to the Krylov subspace (Ritz/Galerkin: CG, CGS)
  - ...
- An important technical question is how to construct the basis for the Krylov subspace if we abandon our simple iterative scheme.

How quickly do the methods converge?

- Most iterative methods perform poorly if applied directly to the original problem $\mathbf{A}\mathbf{x} = \mathbf{b}$.

How quickly do the methods converge?

- Most iterative methods perform poorly if applied directly to the original problem $\mathbf{Ax} = \mathbf{b}$.
- Idea: Solve a ‘preconditioned’ problem

$$\mathcal{P}\mathbf{Ax} = \mathcal{P}\mathbf{b}$$

where $\mathcal{P}$ is chosen such that the iterative method applied to $\mathcal{P}\mathbf{A}$ converges more quickly.

How quickly do the methods converge? (2)

Preconditioned problem

$$\mathcal{P}\mathbf{A}\mathbf{x} = \mathcal{P}\mathbf{b}$$

How quickly do the methods converge? (2)

Preconditioned problem

$$\mathcal{P}\mathbf{A}\mathbf{x} = \mathcal{P}\mathbf{b}$$

- How do we choose/construct $\mathcal{P}$?

How quickly do the methods converge? (2)

Preconditioned problem

$$\mathcal{P}\mathbf{A}\mathbf{x} = \mathcal{P}\mathbf{b}$$

- How do we choose/construct $\mathcal{P}$?
- Simple idea: Use a (cheap!) approximation to the inverse.

$$\mathcal{P} \approx \mathbf{A}^{-1}.$$

How quickly do the methods converge? (2)

Preconditioned problem

$$\mathcal{P}\mathbf{A}\mathbf{x} = \mathcal{P}\mathbf{b}$$

- How do we choose/construct $\mathcal{P}$?
- Simple idea: Use a (cheap!) approximation to the inverse.

$$\mathcal{P} \approx \mathbf{A}^{-1}.$$

- Pointless in its exact form!

How quickly do the methods converge? (2)

Preconditioned problem

$$\mathcal{P}\mathbf{A}\mathbf{x} = \mathcal{P}\mathbf{b}$$

- How do we choose/construct $\mathcal{P}$?
- Simple idea: Use a (cheap!) approximation to the inverse.

$$\mathcal{P} \approx \mathbf{A}^{-1}.$$

- Pointless in its exact form!
- Some popular choices:
  - Diagonal preconditioner (also in block form):

$$\mathcal{P} = (diag\mathbf{A})^{-1}$$

  - Incomplete LU factorisation:

$$\mathcal{P} = (\mathbf{L}_{inc}\mathbf{U}_{inc})^{-1}$$

- In general, the choice of preconditioner is very problem dependent.

Assessing the convergence: The minimal polynomial

- We have already seen that the convergence is related to the eigenvalues of the iteration matrix; $\rho(\mathbf{I} - \mathbf{A}) < 1$ for the Richardson iteration.

Assessing the convergence: The minimal polynomial

- How exactly are the eigenvalues of the matrix related to the convergence of Krylov subspace methods?

Assessing the convergence: The minimal polynomial

- How exactly are the eigenvalues of the matrix related to the convergence of Krylov subspace methods?

- **Definition:** The minimal polynomial of a $N \times N$ nonsingular (diagonalisable) matrix $\mathbf{B}$ with $m \leq N$ distinct eigenvalues λ_j is given by

$$q(t) = \Pi_{j=1}^{m}(t - \lambda_j).$$

Assessing the convergence: The minimal polynomial

- How exactly are the eigenvalues of the matrix related to the convergence of Krylov subspace methods?

- **Definition:** The minimal polynomial of a $N \times N$ nonsingular (diagonalisable) matrix $\mathbf{B}$ with $m \leq N$ distinct eigenvalues λ_j is given by

$$q(t) = \Pi_{j=1}^{m}(t - \lambda_j).$$

- The minimal polynomial has the (obvious) property that

$$q(\mathbf{B}) = \Pi_{j=1}^{m}(\mathbf{B} - \mathbf{I}\lambda_j) = 0.$$

Assessing the convergence: The minimal polynomial

- How exactly are the eigenvalues of the matrix related to the convergence of Krylov subspace methods?

- **Definition:** The minimal polynomial of a $N \times N$ nonsingular (diagonalisable) matrix $\mathbf{B}$ with $m \leq N$ distinct eigenvalues λ_j is given by

$$q(t) = \Pi_{j=1}^{m}(t - \lambda_j).$$

- The minimal polynomial has the (obvious) property that

$$q(\mathbf{B}) = \Pi_{j=1}^{m}(\mathbf{B} - \mathbf{I}\lambda_j) = 0.$$

- Expanding out the polynomial gives

$$q(t) = \Pi_{j=1}^{m}(t - \lambda_j) = \sum_{j=0}^{m} \beta_j t^j$$

for some coefficients β_j.

Assessing the convergence: The minimal polynomial

- How exactly are the eigenvalues of the matrix related to the convergence of Krylov subspace methods?

- **Definition:** The minimal polynomial of a $N \times N$ nonsingular (diagonalisable) matrix $\mathbf{B}$ with $m \leq N$ distinct eigenvalues λ_j is given by

$$q(t) = \Pi_{j=1}^{m}(t - \lambda_j).$$

- The minimal polynomial has the (obvious) property that

$$q(\mathbf{B}) = \Pi_{j=1}^{m}(\mathbf{B} - \mathbf{I}\lambda_j) = 0.$$

- Expanding out the polynomial gives

$$q(t) = \Pi_{j=1}^{m}(t - \lambda_j) = \sum_{j=0}^{m} \beta_j t^j$$

  for some coefficients β_j.

- Applying this to $q(\mathbf{B}) = 0$ gives:

$$0 = \beta_0 \mathbf{I} + \beta_1 \mathbf{B} + ... + \beta_N \mathbf{B}^m$$

Assessing the convergence: The minimal polynomial (2)

$$0 = \beta_0 \mathbf{I} + \beta_1 \mathbf{B} + ... + \beta_m \mathbf{B}^m$$

Assessing the convergence: The minimal polynomial (2)

$$0 = \beta_0 \mathbf{I} + \beta_1 \mathbf{B} + ... + \beta_m \mathbf{B}^m$$

- Multiply through by $\mathbf{B}^{-1}$:

$$\mathbf{B}^{-1} = -\frac{1}{\beta_0} \sum_{j=0}^{m-1} \beta_j \mathbf{B}^j.$$

Assessing the convergence: The minimal polynomial (2)

$$0 = \beta_0 \mathbf{I} + \beta_1 \mathbf{B} + ... + \beta_m \mathbf{B}^m$$

- Multiply through by $\mathbf{B}^{-1}$:

$$\mathbf{B}^{-1} = -\frac{1}{\beta_0} \sum_{j=0}^{m-1} \beta_j \mathbf{B}^j.$$

- So the solution of $\mathbf{Bx} = \mathbf{b}$ can be represented as

$$\mathbf{x} = \mathbf{B}^{-1}\mathbf{b} = -\frac{1}{\beta_0} \sum_{j=0}^{m-1} \beta_j \mathbf{B}^j \mathbf{b}.$$

Assessing the convergence: The minimal polynomial (2)

$$0 = \beta_0 \mathbf{I} + \beta_1 \mathbf{B} + \ldots + \beta_m \mathbf{B}^m$$

- Multiply through by $\mathbf{B}^{-1}$:

$$\mathbf{B}^{-1} = -\frac{1}{\beta_0} \sum_{j=0}^{m-1} \beta_j \mathbf{B}^j.$$

- So the solution of $\mathbf{B}\mathbf{x} = \mathbf{b}$ can be represented as

$$\mathbf{x} = \mathbf{B}^{-1}\mathbf{b} = -\frac{1}{\beta_0} \sum_{j=0}^{m-1} \beta_j \mathbf{B}^j \mathbf{b}.$$

- This shows that

$$\mathbf{x} \in \mathcal{K}_m(\mathbf{B}, \mathbf{b}).$$

Assessing the convergence: The minimal polynomial (3)

- This characterises the solution to $\mathbf{Bx} = \mathbf{b}$ as a member of a Krylov subspace

$$\mathbf{x} \in \mathcal{K}_m(\mathbf{B}, \mathbf{b}),$$

where m is the number of distinct eigenvalues of the matrix $\mathbf{B}$.

- This characterises the solution to $\mathbf{B}\mathbf{x} = \mathbf{b}$ as a member of a Krylov subspace

$$\mathbf{x} \in \mathcal{K}_m(\mathbf{B}, \mathbf{b}),$$

  where m is the number of distinct eigenvalues of the matrix $\mathbf{B}$.

- The fewer distinct eigenvalues the (preconditioned) matrix $\mathbf{B} = \mathcal{P}\mathbf{A}$ has, the lower the degree of the minimal polynomial and the smaller the Krylov subspace, the solution $\mathbf{x}$ lives in.

Assessing the convergence: The minimal polynomial (3)

- This characterises the solution to $\mathbf{Bx} = \mathbf{b}$ as a member of a Krylov subspace

$$\mathbf{x} \in \mathcal{K}_m(\mathbf{B}, \mathbf{b}),$$

  where m is the number of distinct eigenvalues of the matrix $\mathbf{B}$.

- The fewer distinct eigenvalues the (preconditioned) matrix $\mathbf{B} = \mathcal{P}\mathbf{A}$ has, the lower the degree of the minimal polynomial and the smaller the Krylov subspace, the solution $\mathbf{x}$ lives in.

- Recall that during the n-th iteration, a Krylov subspace method will determine the 'best' approximate solution in the Krylov subspace $\mathcal{K}_n(\mathbf{B}, \mathbf{b})$.

Assessing the convergence: The minimal polynomial (3)

- This characterises the solution to $\mathbf{Bx} = \mathbf{b}$ as a member of a Krylov subspace

$$\mathbf{x} \in \mathcal{K}_m(\mathbf{B}, \mathbf{b}),$$

  where m is the number of distinct eigenvalues of the matrix $\mathbf{B}$.

- The fewer distinct eigenvalues the (preconditioned) matrix $\mathbf{B} = \mathcal{P}\mathbf{A}$ has, the lower the degree of the minimal polynomial and the smaller the Krylov subspace, the solution $\mathbf{x}$ lives in.

- Recall that during the n-th iteration, a Krylov subspace method will determine the 'best' approximate solution in the Krylov subspace $\mathcal{K}_n(\mathbf{B}, \mathbf{b})$.

- $\Longrightarrow$ The solution will be found in at most m iterations!

Assessing the convergence: The minimal polynomial (3)

- This characterises the solution to $\mathbf{Bx} = \mathbf{b}$ as a member of a Krylov subspace

$$\mathbf{x} \in \mathcal{K}_m(\mathbf{B}, \mathbf{b}),$$

  where m is the number of distinct eigenvalues of the matrix $\mathbf{B}$.

- The fewer distinct eigenvalues the (preconditioned) matrix $\mathbf{B} = \mathcal{P}\mathbf{A}$ has, the lower the degree of the minimal polynomial and the smaller the Krylov subspace, the solution $\mathbf{x}$ lives in.

- Recall that during the n-th iteration, a Krylov subspace method will determine the 'best' approximate solution in the Krylov subspace $\mathcal{K}_n(\mathbf{B}, \mathbf{b})$.

- $\Longrightarrow$ The solution will be found in at most m iterations!

- This provides an alternative characterisation of a good preconditioner:

> 'A sufficient condition for a good preconditioner is that the preconditioned matrix $\mathbf{B} = \mathcal{P}\mathbf{A}$ has a low degree minimal polynomial.' (Murphy *et al.* 2000)