Disadvantage of this proof:

"continuity" argument does not work for other fields e.g. $\{0,1\} = \mathbb{Z}_2$

Better version of this proof:

triangularise instead of diagonalise.

(not in textbook)

Over $\mathbb{C}$, $\exists$ basis $\mathcal{B} = \{\beta_1, \cdots, \beta_n\}$, such that $[\sigma]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$

(Schur theorem)

so, $\forall k$, $\sigma(\mathrm{Span}\{\beta_1, \cdots, \beta_{k-1}\}) \subseteq \mathrm{Span}\{\beta_1, \cdots, \beta_{k-1}\}$

Furthermore: $\sigma(\beta_k) = *\beta_1 + *\beta_2 + \cdots + *\beta_{k-1} + \lambda_k \beta_k$

$$(\sigma - \lambda_k \iota)(\beta_k) = *\beta_1 + *\beta_2 + \cdots *\beta_{k-1} \in \mathrm{Span}\{\beta_1, \cdots, \beta_{k-1}\}$$

and $(\sigma - \lambda_k \iota)(\mathrm{Span}\{\beta_1, \cdots, \beta_{k-1}\}) \subseteq \mathrm{Span}\{\beta_1, \cdots, \beta_{k-1}\}$

Together: $(\sigma - \lambda_k \iota)(\mathrm{Span}\{\beta_1, \cdots, \beta_k\}) \subseteq \mathrm{Span}\{\beta_1, \cdots, \beta_{k-1}\}$

equivalent: $(\lambda_k \iota - \sigma)(\mathrm{Span}\{\beta_1, \cdots, \beta_k\}) \subseteq \mathrm{Span}\{\beta_1, \cdots, \beta_{k-1}\}$

$$\operatorname{range} \chi_\sigma(\sigma) = \chi_\sigma(\sigma)(\operatorname{Span}\{\beta_1, \cdots, \beta_n\})$$

$$\chi_\sigma(x) = \begin{pmatrix} \lambda_1 - x & & * \\ & \ddots & \\ 0 & & \lambda_n - x \end{pmatrix}$$

$$= (\lambda_1 - x) \ldots (\lambda_n - x)$$

$$= (\lambda_1 \iota - \sigma) \cdots (\lambda_{n-1} \iota - \sigma) \boxed{(\lambda_n \iota - \sigma)(\operatorname{Span}\{\beta_1, \cdots, \beta_n\})}$$

$$\subseteq (\lambda_1 \iota - \sigma) \cdots \boxed{(\lambda_{n-1} \iota - \sigma) \boxed{(\operatorname{Span}\{\beta_1, \cdots, \beta_{n-1}\})}}$$

$$\subseteq (\lambda_1 \iota - \sigma) \cdots (\lambda_{n-2} \iota - \sigma) \boxed{(\operatorname{Span}\{\beta_1, \cdots, \beta_{n-2}\})}$$

$$\vdots$$

$$\subseteq (\lambda_1 \iota - \sigma)(\operatorname{Span}\{\beta_1\}) = \{\vec{0}\}$$

$$\therefore \chi_\sigma(\sigma) = 0 \text{ (zero function)}$$

Back to motivation:

if $\sigma \in L(V,V)$, $\dim V = n$, then $\dim \operatorname{Span}\{\iota, \sigma, \sigma^2, \dots\} \leq n$

$\because \chi_\sigma$ is a degree n polynomial that σ satisfies.

but $\dim \operatorname{Span}\{\iota, \sigma, \sigma^2, \dots\}$ can be smaller, if σ satisfies a polynomial of degree $< n$.

Def 5.2.6: Take $\sigma \in L(V,V)$ $(A = [\sigma]_{\mathscr{A}})$

The minimal polynomial of σ (or of A), written m_σ (or m_A) is the monic polynomial of lowest degree that σ (or A) satisfies

coefficient of highest degree is 1, e.g. $x^2 + ax + b$, $x^3 + ax^2 + bx + c$

It is usually hard to compute m_σ. We focus on its properties.

Facts: i. if m_σ exists, then it is unique.

ii. if $f(\sigma)=0$, then m_σ divides f

(Proof: division algorithm, abstract algebra Th. 5.4)

From ii: if $\dim V < \infty$, then Cayley-Hamilton says that $\boxed{m_\sigma \text{ divides } \chi_\sigma}$

i.e. $\chi_\sigma(x) = m_\sigma(x) f(x)$

(if $\dim V = \infty$, then m_σ might not exist, if σ does not satisfy any polynomials.)

More clearly: if $\chi_\sigma(x) = \pm(x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_k)^{m_k}$ λ_i all distinct

then $m_\sigma(x) = (x-\lambda_1)^{d_1}(x-\lambda_2)^{d_2}\cdots(x-\lambda_k)^{d_k}$

where $\cancel{0} \leq d_i \leq m_i$

e.g. if $\chi_\sigma(x) = -(x-1)^2(x-2)$

then $m_\sigma(x) = (x-1)^{\cancel{0} \text{ or } 1 \text{ or } 2}(x-2)^{\cancel{0} \text{ or } 1}$ ($\cancel{6}$ 2 possibilities)

From Homework: every solution to χ_σ (i.e. every eigenvalue) is a solution to m_σ, so d_i can't be 0.