Fully implicit Runge-Kutta

Ben Southworth, Oliver Krzysik, Will Pazner, Hans de Sterk

June 9, 2020

General Runge-Kutta:

- Multi-stage $\mapsto$ easy implementation.
- Good stability properties.

General Runge-Kutta:

- Multi-stage $\mapsto$ easy implementation.
- Good stability properties.

Fully implicit Runge-Kutta:

- High-order (up to $2s$ for s stages).
- High stage-order (for nonlinear PDEs/DAEs).
- Can compute pairs of stages in parallel.
- Equivalent to DG in time.

Assumptions:

- For Butcher tableaux A_0, $A_0 + A_0^T$ is SPD.
- Linearized operator $\mathcal{L}$ satisfies $\langle \mathcal{L}\mathrm{x}, \mathrm{x} \rangle \leq 0$.

Method of lines for PDEs:

$$M\mathrm{u}'(t) = \mathcal{N}(\mathrm{u}, t) \qquad \text{in } (0, T], \quad \mathrm{u}(0) = \mathrm{u}_0.$$

Runge-Kutta:

$$\mathrm{u}_{n+1} = \mathrm{u}_n + \delta t \sum_{i=1}^{s} b_i \mathrm{k}_i, \qquad \text{where}$$

$$M\mathrm{k}_i = \mathcal{N}\left(\mathrm{u}_n + \delta t \sum_{j=1}^{s} a_{ij}\mathrm{k}_j, t_n + \delta t c_i\right).$$

Method of lines for PDEs:

$$M\mathrm{u}'(t) = \mathcal{N}(\mathrm{u}, t) \qquad \text{in } (0, T], \quad \mathrm{u}(0) = \mathrm{u}_0.$$

Runge-Kutta:

$$\mathrm{u}_{n+1} = \mathrm{u}_n + \delta t \sum_{i=1}^{s} b_i \mathrm{k}_i, \qquad \text{where}$$

$$M\mathrm{k}_i = \mathcal{N}\left(\mathrm{u}_n + \delta t \sum_{j=1}^{s} a_{ij}\mathrm{k}_j, t_n + \delta t c_i\right).$$

Linear(ized) system:

$$\left(\begin{bmatrix} M & & 0 \\ & \ddots & \\ 0 & & M \end{bmatrix} - \delta t \begin{bmatrix} a_{11}\mathcal{L}_1 & \dots & a_{1s}\mathcal{L}_1 \\ \vdots & \ddots & \vdots \\ a_{s1}\mathcal{L}_s & \dots & a_{ss}\mathcal{L}_s \end{bmatrix} \right) \begin{bmatrix} \mathrm{k}_1 \\ \vdots \\ \mathrm{k}_s \end{bmatrix} = \begin{bmatrix} \mathrm{f}_1 \\ \vdots \\ \mathrm{f}_s \end{bmatrix}.$$

Method of lines for PDEs:

$$M\mathrm{u}'(t) = \mathcal{N}(\mathrm{u}, t) \qquad \text{in } (0, T], \quad \mathrm{u}(0) = \mathrm{u}_0.$$

Runge-Kutta:

$$\mathrm{u}_{n+1} = \mathrm{u}_n + \delta t \sum_{i=1}^{s} b_i \mathrm{k}_i, \qquad \text{where}$$

$$M\mathrm{k}_i = \mathcal{N}\left(\mathrm{u}_n + \delta t \sum_{j=1}^{s} a_{ij}\mathrm{k}_j, t_n + \delta t c_i\right).$$

Linear(ized) system:

$$\left(\begin{bmatrix} I & & 0 \\ & \ddots & \\ 0 & & I \end{bmatrix} - \begin{bmatrix} a_{11}\widehat{\mathcal{L}}_1 & \ldots & a_{1s}\widehat{\mathcal{L}}_1 \\ \vdots & \ddots & \vdots \\ a_{s1}\widehat{\mathcal{L}}_s & \ldots & a_{ss}\widehat{\mathcal{L}}_s \end{bmatrix}\right) \begin{bmatrix} \mathrm{k}_1 \\ \vdots \\ \mathrm{k}_s \end{bmatrix} = \begin{bmatrix} \hat{\mathrm{f}}_1 \\ \vdots \\ \hat{\mathrm{f}}_s \end{bmatrix}.$$

Step 1:

$$\left(A_0^{-1} \otimes I - \begin{bmatrix} \widehat{\mathcal{L}}_1 & & \\ & \ddots & \\ & & \widehat{\mathcal{L}}_s \end{bmatrix} \right) (A_0 \otimes I) \begin{bmatrix} \mathsf{k}_1 \\ \vdots \\ \mathsf{k}_s \end{bmatrix} = \begin{bmatrix} \mathsf{f}_1 \\ \vdots \\ \mathsf{f}_s \end{bmatrix}.$$

Step 1:

$$\left(A_0^{-1} \otimes I - \begin{bmatrix} \widehat{\mathcal{L}}_1 & & \\ & \ddots & \\ & & \widehat{\mathcal{L}}_s \end{bmatrix} \right) (A_0 \otimes I) \begin{bmatrix} \mathrm{k}_1 \\ \vdots \\ \mathrm{k}_s \end{bmatrix} = \begin{bmatrix} \mathrm{f}_1 \\ \vdots \\ \mathrm{f}_s \end{bmatrix}.$$

Step 2 (for real Schur decomposition $A_0^{-1} = Q_0 R_0 Q_0^T$):

$$\left(R_0 \otimes I - (Q_0^T \otimes I) \begin{bmatrix} \widehat{\mathcal{L}}_1 & & \\ & \ddots & \\ & & \widehat{\mathcal{L}}_s \end{bmatrix} (Q_0 \otimes I) \right) (R_0^{-1} Q_0^T \otimes I)\mathrm{k} = (Q_0^T \otimes I)\mathrm{f}.$$

Must solve ($\eta > 0$)

$$\begin{bmatrix} \eta I - \widehat{\mathcal{L}}_1 & \phi I \\ -\frac{\beta^2}{\phi} I & \eta I - \widehat{\mathcal{L}}_2 \end{bmatrix},$$

Convergence of Krylov defined by preconditioned Schur complement,

$$\begin{aligned} S &:= \eta I - \widehat{\mathcal{L}}_2 + \beta^2 (\eta I - \widehat{\mathcal{L}}_1)^{-1} \\ &= \left((\eta I - \widehat{\mathcal{L}}_2)(\eta I - \widehat{\mathcal{L}}_1) + \beta^2 I \right) (\eta I - \widehat{\mathcal{L}}_1)^{-1}. \end{aligned}$$

Must solve ($\eta > 0$)

$$\begin{bmatrix} \eta I - \widehat{\mathcal{L}}_1 & \phi I \\ -\frac{\beta^2}{\phi} I & \eta I - \widehat{\mathcal{L}}_2 \end{bmatrix},$$

Convergence of Krylov defined by preconditioned Schur complement,

$$\begin{aligned} S := {} & \eta I - \widehat{\mathcal{L}}_2 + \beta^2 (\eta I - \widehat{\mathcal{L}}_1)^{-1} \\ = {} & \left((\eta I - \widehat{\mathcal{L}}_2)(\eta I - \widehat{\mathcal{L}}_1) + \beta^2 I \right) (\eta I - \widehat{\mathcal{L}}_1)^{-1}. \end{aligned}$$

Consider general preconditioned operator

$$\mathcal{P}_{\delta,\gamma} := \Big[(\eta I - \widehat{\mathcal{L}}_2)(\eta I - \widehat{\mathcal{L}}_1) + \beta^2 I \Big] (\delta I - \widehat{\mathcal{L}}_1)^{-1} (\gamma I - \widehat{\mathcal{L}}_2)^{-1}.$$

Theorem 1 (Optimal preconditioning)

Suppose $\eta > 0$ *and* $W(\widehat{\mathcal{L}}) \leq 0$, *and suppose* $\widehat{\mathcal{L}}$ *is real-valued. Let* $\kappa(\mathcal{P}_{\delta,\gamma})$ *denote the two-norm condition number of* $\mathcal{P}_{\delta,\gamma}$, *for* $\delta, \gamma \in (0, \infty)$, *and define* γ_* *by*

$$\gamma_* := \frac{\eta^2 + \beta^2}{\delta}.$$

Then

$$\kappa(\mathcal{P}_{\delta,\gamma_*}) \leq \frac{1}{2\eta}\left(\delta + \frac{\eta^2 + \beta^2}{\delta}\right).$$

Moreover, (i) the bound is tight in the sense that $\exists\ \widehat{\mathcal{L}}$ *that satisfies the bound exactly, and (ii)* $\gamma = \gamma_*$ *is optimal in the sense that, without further assumptions on* $\widehat{\mathcal{L}}$, γ_* *minimizes a tight upper bound on* $\kappa(\mathcal{P}_{\delta,\gamma})$ *over all* $\gamma \in (0, \infty)$.

Corollary 2 (Optimal preconditioning with $\gamma = \delta$)

A tight upper bound on the condition number of $\mathcal{P}_{\delta,\gamma}$ is minimized over $\delta, \gamma \in (0, \infty)$ with $\delta = \gamma = \gamma_* = \sqrt{\eta^2 + \beta^2}$. Furthermore, the condition number of $\mathcal{P}_{\gamma_*}$ is tightly bounded via

$$\kappa(\mathcal{P}_{\gamma_*}) \leq \sqrt{1 + \frac{\beta^2}{\eta^2}}.$$

Stages	2	3		4		5		
	$\lambda^{\pm}_{1,2}$	λ_1	$\lambda^{\pm}_{2,3}$	$\lambda^{\pm}_{1,2}$	$\lambda^{\pm}_{3,4}$	λ_1	$\lambda^{\pm}_{2,3}$	$\lambda^{\pm}_{4,5}$
Gauss	1.15	1.00	1.38	1.61	1.04	1.00	1.83	1.13
Radau	1.22	1.00	1.51	1.79	1.05	1.00	2.05	1.15
Lobatto	1.41	1.00	1.79	2.12	1.06	1.00	2.42	1.17

Table: Bounds on $\kappa(\mathcal{P}_{\gamma_*})$ for Gauss, Radau IIA, and Lobatto IIIC.

Corollary 3 (Optimal preconditioning with $\gamma = \delta$)

A tight upper bound on the condition number of $\mathcal{P}_{\eta,\gamma}$ is minimized over $\gamma \in (0, \infty)$ with $\delta = \gamma = \gamma_* = \eta + \frac{\beta^2}{\eta}$. Furthermore, the condition number of $\mathcal{P}_{\gamma_*}$ is tightly bounded via

$$\kappa(\mathcal{P}_{\gamma_*}) \le 1 + \frac{1}{2}\frac{\beta^2}{\eta^2}.$$

Stages	2	3		4		5		
	$\lambda_{1,2}^{\pm}$	λ_1	$\lambda_{2,3}^{\pm}$	$\lambda_{1,2}^{\pm}$	$\lambda_{3,4}^{\pm}$	λ_1	$\lambda_{2,3}^{\pm}$	$\lambda_{4,5}^{\pm}$
Gauss	1.17	1.00	1.46	1.80	1.05	1.00	2.18	1.14
Radau IIA	1.25	1.00	1.65	2.11	1.06	1.00	2.60	1.16
Lobatto IIIC	1.50	1.00	2.11	2.76	1.07	1.00	3.44	1.19

Table: Bounds on $\kappa(\mathcal{P}_{\gamma_*})$ for Gauss, Radau IIA, and Lobatto IIIC.