

ON THE ITERATIVE SOLUTION OF SADDLE POINT PROBLEMS

Philippe Devloo Giovane Avancini Jeferson Fernandes Sonia Gomes





Outline

- Two different preconditioners are presented
- A spectrally efficient preconditioner for the Darcy equation
- 2. A preconditioner for incompressible elasticity or the Stokes equation





Part 1 An efficient preconditioner for mixed finite element approximations





Motivation

Mixed finite element simulations are applicable to reservoir simulation: they allow using unstructured meshes with a higher precision when compared to FV.

The goals of this work are:

- Develop a fast prototype solver;
- Derive low cost and efficient preconditioners for high order mixed finite element computations and reservoir simulations;
- Obtain local conservation for the approximate solutions at each iteration;







Methodology

First step: development of low cost div-free and div-constant H(div) spaces



LABMEC

Reduction of **55%-65%** in # DoF for the same result.

Reduction of **45%-50%** in # DoF for the same result.



Methodology

LABMEC

Application: flow in heterogeneous porous media.



P. R. B. Devloo et al. "An efficient construction of divergence-free spaces in the context of exact finite element de Rham sequences". *Computer Methods in Applied Mechanics and Engineering* (2022).



Methodology

Second step: move forward and solve the reduced problem in a more efficient way.

$$\begin{bmatrix} A & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} \sigma \\ u \end{bmatrix} = \begin{bmatrix} f \\ 0 \end{bmatrix} \Rightarrow \mathbf{K}\mathbf{u} = \mathbf{F}$$

Direct solvers

Involves the inversion (or decomposition)

of matrix **K**;

In practice, its efficiency is limited to the matrix dimensions and sparsity.

Iterative solvers

Do not require the inversion of matrix **K** but uses projections to correct an initial guess solution.

Matrix **K** needs to be properly conditioned;

Suitable for large scale problems.





Model problem - Darcy flow Semi-hybrid form

Now the formulation is stated as: find $\tilde{\boldsymbol{\sigma}} = \hat{\boldsymbol{\sigma}} + \boldsymbol{\sigma}^0 \in \hat{\mathbf{V}}_{\eta^0} \oplus \mathbf{V}_{\eta}^0$, with $\hat{\boldsymbol{\sigma}} \cdot \mathbf{n}|_{\partial\Omega_N} = \bar{g}$, $\boldsymbol{\sigma}^0 \cdot \mathbf{n}|_{\partial\Omega_N} = g^{\perp}$ $\tilde{u} \in W_{\eta^0}$ and $\tilde{\lambda} \in \Lambda_{\eta^0}$ such that

$$\begin{split} \int_{\Omega} \mathbb{K}^{-1} \hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{v}} \, d\mathbf{x} + \int_{\Omega} \mathbb{K}^{-1} \boldsymbol{\sigma}^{0} \cdot \hat{\mathbf{v}} \, d\mathbf{x} - \sum_{K \in \mathcal{T}_{h}} \int_{K} \tilde{u} \, \nabla \cdot \hat{\mathbf{v}} \, d\mathbf{x} + \sum_{K \in \mathcal{T}_{h}} \int_{\partial K \setminus \partial \Omega} \tilde{\lambda} \, \hat{\mathbf{v}} \cdot \mathbf{n}^{K} ds &= -\int_{\partial \Omega_{D}} u_{D}(\hat{\mathbf{v}} \cdot \mathbf{n}) ds, \\ \int_{\Omega} \mathbb{K}^{-1} \hat{\boldsymbol{\sigma}} \cdot \mathbf{v} \, d\mathbf{x} + \int_{\Omega} \mathbb{K}^{-1} \boldsymbol{\sigma}^{0} \cdot \mathbf{v} \, d\mathbf{x} &= -\int_{\partial \Omega_{D}} u_{D}(\mathbf{v} \cdot \mathbf{n}) ds, \\ \sum_{K \in \mathcal{T}_{h}} \int_{K} \nabla \cdot \hat{\boldsymbol{\sigma}} \, w d\mathbf{x} &= \int_{\Omega} f w d\mathbf{x}, \\ \sum_{K \in \mathcal{T}_{h}} \int_{\partial K \setminus \partial \Omega} \mu \, \hat{\boldsymbol{\sigma}} \cdot \mathbf{n}^{K} ds &= 0, \end{split}$$





Model problem - Darcy flow Iterative method

Static condensation can be applied to the problem:

$$\begin{bmatrix} \mathbf{M}_{\mu\mu} & \mathbf{M}_{\mu\mathbf{v}^{e}} \\ \mathbf{M}_{\mathbf{v}^{e}\mu} & \mathbf{M}_{\mathbf{v}^{e}\mathbf{v}^{e}} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda} \\ \mathbf{S}^{e} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{\mu} \\ \mathbf{G}_{\mathbf{v}^{e}} \end{bmatrix} \Longrightarrow \begin{bmatrix} \mathbf{M}_{\mu\mu} \mathbf{\Lambda} = \mathbf{G}_{\mu} - \mathbf{M}_{\mu\mathbf{v}^{e}} \mathbf{S}^{e}. \tag{1}$$

$$\begin{pmatrix} \mathbf{M}_{\mathbf{v}^{e}\mathbf{v}^{e}} - \mathbf{M}_{\mathbf{v}^{e}\mu} \mathbf{M}_{\mu\mu}^{-1} \mathbf{M}_{\mu\mathbf{v}^{e}} \end{pmatrix} \mathbf{S}^{e} = \mathbf{G}_{\mathbf{v}^{e}} - \mathbf{M}_{\mathbf{v}^{e}\mu} \mathbf{M}_{\mu\mu}^{-1} \mathbf{G}_{\mu}. \tag{2}$$

ALGORITHM

- Decompose matrix M_{µµ}, which has dimension independent of the polynomial degree, as it corresponds to the constant flux approximations, and is equal to the corresponding FV problem.
- Solve iteratively problem (2) applying the conjugate gradient method and a block diagonal preconditioner;
- Once problem (1)-(2) reaches convergence, the condensed variables can be explicitly recovered.
- Remark: A conservative solution can be obtained at each iteration;





Numerical Results Problem setup

- A tolerance of $\varepsilon = 10^{-10}$ (Euclidean norm) in the iterative process was considerd;
- The permeability tensor is equal to the identity, namely K = I
- I Numerical tests were performed with quadrilateral (in 2D) and hexahedral (in 3D) meshes for several configurations, with both h and p refinement.
- The following exact solutions were adopted:

2D problem $u = e^{\pi x} \sin(\pi y)$ $u = \frac{\sin(\pi x) \sin(\pi y) \sinh(\sqrt{2}\pi z)}{\sinh(\sqrt{2}\pi)}$





2D case - h refinement



k: polynomial order

LABMEC

blue : size of the constant pressure matrix



2D case - p refinement

LABMEC





3D case - h refinement



k: polynomial order

LABMEC

blue : size of the constant pressure matrix



3D case - p refinement

LABMEC





Efficient preconditioner obtained, with low number of iterations;

The number of iterations doesn't increase mesh refinement;

The number of iterations increases in a log scale with polynomial degree;

Theoretical results for the matrix condition number were obtained with optimal rate;

Good results for both 2D and 3D approaches;

High reduction in the number of equations of the global system.

A conservative solution can be obtained at each iteration;

Conclusions





Part 2 An efficient preconditioner for mixed finite element approximations of Stokes equations





Outline

- Challenges in inverting saddle point problems
- Problem statement
- A preconditioner introducing a small compressibility
- A positive definite preconditioner
- Numerical results
- Conclusions





Motivation

 A linear system of equations modeling a saddle point problem has a structure as

$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{0} \end{bmatrix}$

- Depending on the ordering of the equations, its decomposition will require pivoting equations
- If small numerical perturbations occur during decomposition, the linear solver may not pivot where necessary and give wrong results





Applications

The iterative scheme is tested for two kind of boundary value problems:

1) Darcy equations

> 2) Stokes or incompressible elasticity equations





Darcy problem statement

$$\sigma = -\mathcal{K}\nabla p, \quad \text{in} \quad \Omega,$$

 $\nabla \cdot \sigma = f, \quad \text{in} \quad \Omega,$
 $p = p_D, \quad \text{on} \quad \partial \Omega_D,$
 $\sigma \cdot n = g, \quad \text{on} \quad \partial \Omega_N,$

 The finite element approximation using De Rham compatible flux/pressure spaces is

find $\sigma^{\gamma} \in \mathbf{V}^{\gamma}$ and $p^{\gamma} \in W^{\gamma}$ such that for all $\mathbf{w}^{\gamma}_{\sigma} \in \mathbf{V}^{\gamma}$ and $w^{\gamma}_{p} \in W^{\gamma}$:

$$\sum_{\Omega_e \in \mathcal{T}} \left(\int_{\Omega_e} \mathcal{K}^{-1} \boldsymbol{\sigma}^{\gamma} \cdot \mathbf{w}_{\boldsymbol{\sigma}}^{\gamma} \, d\Omega - \int_{\Omega_e} p^{\gamma} \, \nabla \cdot \mathbf{w}_{\boldsymbol{\sigma}}^{\gamma} \, d\Omega \right) = -\sum_{\Omega_e \in \partial\Omega_D} \int_{\partial\Omega_e} p_D \left(\mathbf{w}_{\boldsymbol{\sigma}}^{\gamma} \cdot \mathbf{n} \right) d\partial\Omega,$$
$$\sum_{\Omega_e \in \mathcal{T}} \int_{\Omega_e} \nabla \cdot \boldsymbol{\sigma}^{\gamma} \, w_p^{\gamma} \, d\Omega = \sum_{\Omega_e \in \mathcal{T}} \int_{\Omega_e} f w_p^{\gamma} \, d\Omega.$$





Darcy formulation properties

- For each element, the flux shape functions can be partitioned
 - Internal fluxes
 - Boundary fluxes
- The pressure functions can be partitioned
 - Pressures with zero average
 - Constant pressure
- The internal fluxes and pressures with zero average value can be statically condensed at the element level

$$[K]_{el} = \begin{bmatrix} K_{\sigma_n \sigma_n} & K_{\sigma_n \bar{p}} \\ K_{\bar{p} \sigma_n} & 0 \end{bmatrix}$$



Stokes problem statement

$$egin{aligned} -
abla \cdot oldsymbol{\sigma} &= oldsymbol{f} & ext{in} & \Omega, \ -
abla \cdot \mathbf{u} &= \mathbf{0} & ext{in} & \Omega, \ \mathbf{u} &= \mathbf{u}_D & ext{on} & \partial\Omega_D, \ oldsymbol{\sigma} &= oldsymbol{\sigma}_N & ext{on} & \partial\Omega_N, \end{aligned}$$

• A double-hybrid finite element formulation with compatible De Rham spaces is developed, resulting in

Find $\mathbf{u}, p, \boldsymbol{\lambda}^t, \mathbf{u}^t \in \vec{\mathbb{V}}_D \times \mathbb{Q}_d \times \boldsymbol{\Lambda}^t_* \times \boldsymbol{\mathcal{L}}_D^t$ such that for all $\mathbf{v}, q, \boldsymbol{\eta}^t, \mathbf{v}^t \in \vec{\mathbb{V}}_0 \times \mathbb{Q}_d \times \boldsymbol{\Lambda}^t_* \times \boldsymbol{\mathcal{L}}_D^t$, the following equations are satisfied

$$\sum_{\Omega_{e} \in \mathcal{T}} \int_{\Omega_{e}} 2\mu \boldsymbol{\varepsilon}(\mathbf{v}) \cdot \boldsymbol{\varepsilon}(\mathbf{u}) d\Omega_{e} - \sum_{\Omega_{e} \in \mathcal{T}} \int_{\Omega_{e}} p(\nabla \cdot \mathbf{v}) d\Omega_{e} - \sum_{\partial \Omega_{e} \in \partial \mathcal{T}} \int_{\partial \Omega_{e}} \boldsymbol{\lambda}^{t} \cdot \mathbf{v} d\partial \Omega_{e} = \sum_{\Omega_{e} \in \mathcal{T}} \int_{\Omega_{e}} \mathbf{v} \cdot \mathbf{f} d\Omega_{e} + \sum_{E \in \partial \Omega_{N}} \int_{E} \boldsymbol{\sigma}_{N}^{n} (\mathbf{v} \cdot \mathbf{n}) d\partial \Omega_{e}$$
$$\sum_{\Omega_{e} \in \mathcal{T}} - \int_{\Omega_{e}} q(\nabla \cdot \vec{u}) d\Omega_{e} = 0$$
$$\sum_{\partial \Omega_{e} \in \partial \mathcal{T}} - \int_{\partial \Omega_{e}} \mathbf{u} \cdot \boldsymbol{\eta}^{t} d\partial \Omega_{e} + \sum_{E \in \varepsilon_{0}} \int_{E} [\![\boldsymbol{\eta}^{t}]\!] \cdot \mathbf{u}^{t} d\partial \Omega_{e} = 0$$
$$\sum_{\partial \Omega_{e} \in \partial \mathcal{T}} \int_{E} [\![\boldsymbol{\lambda}^{t}]\!] \cdot \mathbf{v}^{t} d\partial \Omega_{e} = \sum_{E \in \partial \Omega_{N}} \int_{E} \mathbf{v}^{t} \cdot (\boldsymbol{\lambda}^{t} - \boldsymbol{\sigma}_{N}^{t}) d\partial \Omega_{e}$$



Stokes formulation properties

- For each element, the velocity shape functions can be partitioned
 - Internal velocities
 - Boundary normal velocities
- The pressure functions can be partitioned
 - Pressures with zero average
 - Constant pressure
- The internal velocities pressures with zero average value and tangential stresses can be statically condensed at the element level

$$[K]_{el} = \begin{bmatrix} K_{\mathbf{u}\mathbf{u}^t} & K_{\mathbf{u}p} \\ K_{\mathbf{u}p}^T & 0 \end{bmatrix}$$



Global system of equations

Grouping the fluxes first and pressures second

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{0} \end{bmatrix} \begin{pmatrix} \boldsymbol{\sigma} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{f}_{\boldsymbol{\sigma}} \\ \mathbf{f}_p \end{pmatrix}$$

- The matrix A is positive definite
- Each element contributes to one equation in B



Preconditioner

 A preconditioner is proposed adding a small compressibility to the global system of eqs

$$\begin{bmatrix} \mathbf{G} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{0} \end{bmatrix} \implies \tilde{\mathbf{G}} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & -\mathbf{C} \end{bmatrix}$$

- Where C is a positive definite matrix representing an "artificial" compressibility
- The pressure equations can now be statically condensed

 $ar{\mathbf{G}} = \mathbf{A} + \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T$ $ar{\mathbf{f}} = \mathbf{f}_{\boldsymbol{\sigma}} + \mathbf{B}\mathbf{C}^{-1}\mathbf{f}_p$





Iterative method

 $\bar{\mathbf{G}} = \mathbf{A} + \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T \qquad \quad \bar{\mathbf{f}} = \mathbf{f}_{\boldsymbol{\sigma}} + \mathbf{B}\mathbf{C}^{-1}\mathbf{f}_p$

Compute an initial solution for flux and pressure

$$\boldsymbol{\sigma}^{0} = \bar{\mathbf{G}}^{-1}\bar{\mathbf{f}}$$
 $\mathbf{p}^{0} = C^{-1}\left(\mathbf{B}^{T}\boldsymbol{\sigma}^{0} - \mathbf{f}_{p}\right)$

Compute the residual $\mathbf{r}^{k} = -\mathbf{B}\mathbf{C}^{-1}(\mathbf{B}^{T}\boldsymbol{\sigma}^{k} - \mathbf{f}_{p})$ The original constraint
The original constraint
Solve the system to obtain the flux correction $\Delta \boldsymbol{\sigma}^{k} = \bar{\mathbf{G}}^{-1}\mathbf{r}^{k}$ $\boldsymbol{\sigma}^{k+1} = \boldsymbol{\sigma}^{k} + \Delta \boldsymbol{\sigma}^{k}$

- Update the pressure explicitly $\Delta \mathbf{p}^k = C^{-1} \mathbf{B}^T \Delta \boldsymbol{\sigma}^k \qquad \qquad \mathbf{p}^{k+1} = \mathbf{p}^k + \Delta \mathbf{p}^k$
- Notice that the flux update does not depend on the pressure!





Convergence of the method

$$[\mathbf{G}] = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{G}} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & -\mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{G} \end{bmatrix} [\mathbf{U}] = [\mathbf{F}]$$

$$\begin{bmatrix} \mathbf{U} \end{bmatrix}^{k+1} - \begin{bmatrix} \mathbf{U} \end{bmatrix}^{k} = \begin{bmatrix} \bar{\mathbf{G}} \end{bmatrix}^{-1} \left(\begin{bmatrix} \mathbf{F} \end{bmatrix} - \begin{bmatrix} \mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{U} \end{bmatrix}^{k} \right)$$
$$\begin{bmatrix} \mathbf{U} \end{bmatrix}^{k+2} - \begin{bmatrix} \mathbf{U} \end{bmatrix}^{k+1} = \begin{bmatrix} \bar{\mathbf{G}} \end{bmatrix}^{-1} \left(\begin{bmatrix} \mathbf{F} \end{bmatrix} - \begin{bmatrix} \mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{U} \end{bmatrix}^{k+1} \right)$$
$$\begin{bmatrix} \mathbf{U} \end{bmatrix}^{k+1} - \mathbf{\Delta} \begin{bmatrix} \mathbf{U} \end{bmatrix}^{k} = \begin{bmatrix} \bar{\mathbf{G}} \end{bmatrix}^{-1} \left(-\begin{bmatrix} \mathbf{G} \end{bmatrix} \Delta \begin{bmatrix} \mathbf{U} \end{bmatrix}^{k} \right)$$

$$\Delta \left[\mathbf{U} \right]^{k+1} = \left(\left[I \right] - \left[\bar{\mathbf{G}} \right]^{-1} \left[\mathbf{G} \right] \right) \Delta \left[\mathbf{U} \right]^{k}$$





$$\begin{split} \mathbf{\bar{G}} &= \mathbf{G} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{C} \end{bmatrix} \qquad \begin{array}{c} \text{Taylor series expansion} \\ \mathbf{\bar{G}}^{-1} &= \mathbf{G}^{-1} - \mathbf{G}^{-1} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{C} \end{bmatrix} \mathbf{G}^{-1} + O(\mathbf{C}) \\ (\mathbf{I} - \mathbf{\bar{G}}^{-1}\mathbf{G}) &\simeq \mathbf{G}^{-1} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{C} \end{bmatrix} \qquad \begin{array}{c} \Delta[\mathbf{U}]^{k+1} = \left([I] - [\mathbf{\bar{G}}]^{-1}[\mathbf{G}] \right) \Delta[\mathbf{U}]^k \end{split}$$

- If C is suficiently small, the method will converge
- QUESTION: How to define matrix C?

$$[C] = \alpha h_e[I]$$

Artificial compressibility parameter





Where are we different from Mardal/Winter?

CONSTRUCTION OF PRECONDITIONERS BY MAPPING PROPERTIES FOR SYSTEMS OF PARTIAL DIFFERENTIAL EQUATIONS

KENT-ANDRE MARDAL¹ AND RAGNAR WINTHER²

$$\mathcal{A} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} -\Delta & -\operatorname{grad} \\ \operatorname{div} & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}. \qquad \qquad \mathcal{B} = \begin{pmatrix} -\Delta^{-1} & 0 \\ 0 & I \end{pmatrix}$$
$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{0} \end{bmatrix} \begin{pmatrix} \boldsymbol{\sigma} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{f}_{\boldsymbol{\sigma}} \\ \mathbf{f}_{p} \end{pmatrix} \qquad \qquad \tilde{\mathbf{G}} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & -\mathbf{C} \end{bmatrix}$$

We just found a "clever" way of efficiently inverting ${ar G}$





Darcy Numerical results

• Given a 3D Darcy problem on a unit cube

 $\mathcal{K} = \mathbf{I}_3.$

$$p(x, y, z) = \frac{\sin(\pi x)\sin(\pi y)\sinh(\sqrt{2\pi}z)}{\sinh(\sqrt{2\pi}z)}$$









Convergence behavior

Convergence is very smooth Optimum Alpha between 10⁻³ and 10⁻²

If alpha is too small roundoff error will limit the convergence







LABMEC

Acceleration

Compare direct solver with iterative solver We gain in cpu time We gain in stability



Stokes Numerical results

Poiseuille flow in a rectangular domain

[0,2]x[-1,1]x[-1,1] $\mu = 1.0$

$$u_x(y) = \frac{\nabla p}{2\mu}(1-y^2)$$







Convergence behavior

Convergence is very smooth Optimum Alpha between 10⁻⁴ and 10⁻³







LABMEC

Compare direct solver with iterative solver We gain in cpu time We gain in stability



Conclusion Stokes/Darcy

- An iterative method was presented for solving saddle point problems using a positive definite preconditioner
- The system of the preconditioner is smaller than the global problem
- The iterative method was up to 30% and 40% faster than the direct method for Darcy and Stokes, respectively





Acknowledgement

- Fapesp/Equinor grant 2017/15736-3
- CNPq grant 305823-2017-5
- TotalEnergies grant TEPBr/RD21-14 SAP n° 4600003236

