

# Some more comments on the normal equations: With a focus on discretisation of partial differential equations

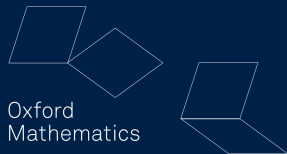


Mathematical  
Institute

L. Lazzarino, Y. Nakatsukasa, Umberto Zerbinati\*

*\*Mathematical Institute – University of Oxford*

Due Giorni Di Algebra Lineare Numerica, 20th January 2025



# THE NORMAL EQUATION

Let us consider the following linear system of equations

$$\underline{\underline{A}}\underline{\underline{x}} = \underline{\underline{b}}, \quad \underline{\underline{A}} \in \mathbb{R}^{n \times n}, \quad \underline{\underline{x}}, \underline{\underline{b}} \in \mathbb{R}^n.$$

$$A^T \neq A$$

In order to solve the system, we can consider the normal equation, i.e.

$$B := \underline{\underline{A}}^T \underline{\underline{A}} \underline{\underline{x}} = \underline{\underline{A}}^T \underline{\underline{b}}$$



*SIMAX* Vol. 13, Iss. 3, 1992 (N. M. Nachtigal, S. C. Reddy, L. N. Trefethen),

L. N. Trefethen and D. Bau, III, *Numerical Linear Algebra*, 1997, *SIAM*.

# THE NORMAL EQUATION

Let us consider the following linear system of equations

$$\underline{\underline{A}}\underline{\underline{x}} = \underline{\underline{b}}, \quad \underline{\underline{A}} \in \mathbb{R}^{n \times n}, \quad \underline{\underline{x}}, \underline{\underline{b}} \in \mathbb{R}^n.$$

$$A^T \neq A$$

In order to solve the system, we can consider the normal equation, i.e.

$$B := \underline{\underline{A}}^T \underline{\underline{A}} \underline{\underline{x}} = \underline{\underline{A}}^T \underline{\underline{b}}$$

► How to **quickly** access  $\underline{\underline{A}}^T$  and  $\underline{\underline{B}}$  ?



*SIMAX* Vol. 13, Iss. 3, 1992 (N. M. Nachtigal, S. C. Reddy, L. N. Trefethen),

L. N. Trefethen and D. Bau, III, *Numerical Linear Algebra*, 1997, *SIAM*.

# THE NORMAL EQUATION

Let us consider the following linear system of equations

$$\underline{\underline{A}}\underline{\underline{x}} = \underline{\underline{b}}, \quad \underline{\underline{A}} \in \mathbb{R}^{n \times n}, \quad \underline{\underline{x}}, \underline{\underline{b}} \in \mathbb{R}^n.$$

$$A^T \neq A$$

In order to solve the system, we can consider the normal equation, i.e.

$$B := \underline{\underline{A}}^T \underline{\underline{A}} \underline{\underline{x}} = \underline{\underline{A}}^T \underline{\underline{b}}$$

► How to **quickly** access  $\underline{\underline{A}}^T$  and  $\underline{\underline{B}}$  ?



*SIMAX* Vol. 13, Iss. 3, 1992 (N. M. Nachtigal, S. C. Reddy, L. N. Trefethen),

L. N. Trefethen and D. Bau, III, Numerical Linear Algebra, 1997, *SIAM*.

► Unfortunately the condition number of  $\underline{\underline{A}}^T \underline{\underline{A}}$  is the square of the condition number of  $\underline{\underline{A}}$ .

# THE NORMAL EQUATION

Let us consider the following linear system of equations

$$\underline{\underline{A}}\underline{\underline{x}} = \underline{\underline{b}}, \quad \underline{\underline{A}} \in \mathbb{R}^{n \times n}, \quad \underline{\underline{x}}, \underline{\underline{b}} \in \mathbb{R}^n.$$

$$A^T \neq A$$

In order to solve the system, we can consider the normal equation, i.e.

$$B := \underline{\underline{A}}^T \underline{\underline{A}} \underline{\underline{x}} = \underline{\underline{A}}^T \underline{\underline{b}}$$

► How to **quickly** access  $\underline{\underline{A}}^T$  and  $\underline{\underline{B}}$  ?




*SIMAX* Vol. 13, Iss. 3, 1992 (N. M. Nachtigal, S. C. Reddy, L. N. Trefethen),

L. N. Trefethen and D. Bau, III, Numerical Linear Algebra, 1997, *SIAM*.

- Unfortunately the condition number of  $\underline{\underline{A}}^T \underline{\underline{A}}$  is the square of the condition number of  $\underline{\underline{A}}$ .
- We now have a symmetric positive definite system, that can be solved using CG (**CGNE**).


# HOW CAN WE PRECONDITION THE NORMAL EQUATIONS?


*SIREV* Vol. 64, Iss. 3, 2022 (A. Wathen),

## Good preconditioners – Classical Definition

$\underline{\underline{P}}$  is a good preconditioner if  $\underline{\underline{P}}^{-1}\underline{\underline{A}}$  has clustered eigenvalues.

# HOW CAN WE PRECONDITION THE NORMAL EQUATIONS?


*SIREV* Vol. 64, Iss. 3, 2022 (A. Wathen),

## Good preconditioners – Classical Definition

$\underline{\underline{P}}$  is a good preconditioner if  $\underline{\underline{P}}^{-1}\underline{\underline{A}}$  has clustered eigenvalues.

Unfortunately given a good preconditioner  $\underline{\underline{P}}$  for  $\underline{\underline{A}}$  we might not have good preconditioner  $\underline{\underline{G}} := \underline{\underline{P}}^T \underline{\underline{P}}$  for  $\underline{\underline{A}}^T \underline{\underline{A}}$ .

# HOW CAN WE PRECONDITION THE NORMAL EQUATIONS?

📖 *SIREV* Vol. 64, Iss. 3, 2022 (A. Wathen),

## Good preconditioners – Classical Definition

$\underline{\underline{P}}$  is a good preconditioner if  $\underline{\underline{P}}^{-1}\underline{\underline{A}}$  has clustered eigenvalues.

Unfortunately given a good preconditioner  $\underline{\underline{P}}$  for  $\underline{\underline{A}}$  we might not have good preconditioner  $\underline{\underline{G}} := \underline{\underline{P}}^T \underline{\underline{P}}$  for  $\underline{\underline{A}}^T \underline{\underline{A}}$ .

$$A = \begin{bmatrix} b_0 & & \\ & \ddots & \\ & & b_{n-1} \end{bmatrix}, \quad P = \begin{bmatrix} & & b_0 \\ & \ddots & \\ b_{n-1} & & \end{bmatrix}.$$



# HOW CAN WE PRECONDITION THE NORMAL EQUATIONS?

📖 *SIREV* Vol. 64, Iss. 3, 2022 (A. Wathen),

## Good preconditioners – Classical Definition

$\underline{\underline{P}}$  is a good preconditioner if  $\underline{\underline{P}}^{-1}\underline{\underline{A}}$  has clustered eigenvalues.

Unfortunately given a good preconditioner  $\underline{\underline{P}}$  for  $\underline{\underline{A}}$  we might not have good preconditioner  $\underline{\underline{G}} := \underline{\underline{P}}^T \underline{\underline{P}}$  for  $\underline{\underline{A}}^T \underline{\underline{A}}$ .

$$P^{-1}A = \begin{bmatrix} & & 1 \\ & \ddots & \\ 1 & & \end{bmatrix}, \quad G^{-1}B = \begin{bmatrix} (b_0/b_{n-1})^2 & & \\ & \ddots & \\ & & (b_{n-1}/b_0)^2 \end{bmatrix}.$$

# HOW CAN WE PRECONDITION THE NORMAL EQUATIONS?



*SIREV* Vol. 64, Iss. 3, 2022 (A. Wathen),  
*QJRM* Vol. 64, Iss. 114, 2018 (S. Gratton, Et Al.).

## Gratton–Gürol–Simon–Toint

If the matrix  $P$  is such that  $\|I - AP^{-1}\|_2 \leq \sqrt{2} - 1 - \delta$ , then  

$$\Lambda(G^{-1}B) \subset (\sqrt{2}\delta + \delta^2, 2 - \sqrt{2}\delta - \delta^2).$$

## HOW CAN WE PRECONDITION THE NORMAL EQUATIONS?



*SIREV* Vol. 64, Iss. 3, 2022 (A. Wathen),  
*QJRM* Vol. 64, Iss. 114, 2018 (S. Gratton, Et Al.).

### Gratton–Gürol–Simon–Toint

If the matrix  $P$  is such that  $\|I - AP^{-1}\|_2 \leq \sqrt{2} - 1 - \delta$ , then  
 $\Lambda(G^{-1}B) \subset (\sqrt{2}\delta + \delta^2, 2 - \sqrt{2}\delta - \delta^2)$ .

We consider the matrix  $T := I - AP^{-1}$ , and expand  $G^{-1}B$  as

$$G^{-1}B = P^{-1}P^{-T}A^T A \sim P^{-T}A^T AP^{-1} = I - T - T^T + T^T T.$$

Since  $\Lambda(G^{-1}B) \subset [-\|G^{-1}B\|_2, \|G^{-1}B\|_2]$ , we can easily see that

$$-1 - 2\|T\|_2 - \|T\|_2^2 \leq \lambda \leq 1 + 2\|T\|_2 + \|T\|_2^2.$$

Substituting  $\|I - AP^{-1}\|_2 \leq \sqrt{2} - 1 - \delta$  we obtained the desired result.

## CROSS PRECONDITIONING

We would like to give a different intuition of good preconditioners for normal equations. To this aim we consider the previously observed similarity,

$$G^{-1}B = P^{-1}P^{-T}A^TA \sim P^{-T}A^TAP^{-1} = (AP^{-1})^T(AP^{-1}).$$

Hence, the closer the matrix  $AP^{-1}$  is to an orthogonal matrix, the closer  $G^{-1}B$  is to the identity matrix.

## CROSS PRECONDITIONING

We would like to give a different intuition of good preconditioners for normal equations. To this aim we consider the previously observed similarity,

$$G^{-1}B = P^{-1}P^{-T}A^TA \sim P^{-T}A^TAP^{-1} = (AP^{-1})^T(AP^{-1}).$$

Hence, the closer the matrix  $AP^{-1}$  is to an orthogonal matrix, the closer  $G^{-1}B$  is to the identity matrix.

### Cross preconditioning

We say that the preconditioner  $P$  is a good **left** preconditioner for the normal equations if it is a good **right** preconditioner for  $\underline{\underline{A}}$ , in the sense that  $\underline{\underline{AP}}^{-1}$  has **clustered singular values**.

## ADVECTION DIFFUSION ODE – CROSS PRECONDITIONING

We consider the classical advection-diffusion ODE in one dimension, i.e.

$$\begin{aligned}
 & -\nu \ddot{u} + \beta \dot{u} = f \text{ in } (a, b) \subset \mathbb{R}, \\
 & u(a) = 0, \quad u(b) = 1, \quad \nu, \beta \in \mathbb{R}_{\geq 0}.
 \end{aligned}$$



R. J. LeVeque, Finite Difference Methods for Ordinary and Partial Differential Equations, 2007 ,SIAM.

## ADVECTION DIFFUSION ODE – CROSS PRECONDITIONING

We consider the classical advection-diffusion ODE in one dimension, i.e.

$$\begin{aligned} -\nu \ddot{u} + \beta \dot{u} &= f \text{ in } (a, b) \subset \mathbb{R}, \\ u(a) &= 0, \quad u(b) = 1, \quad \nu, \beta \in \mathbb{R}_{\geq 0}. \end{aligned}$$

For the moment we will consider neither diffusion nor advection-dominated regimes, i.e.  $\nu \approx \beta$ , and discretisation over an equi-spaced mesh of step-size  $h$ . Such a discretisation results in the matrix

$$\underline{\underline{A}} = \text{tridiag} \left( -\frac{\nu}{h^2} - \frac{\beta}{2h}, \frac{2\nu}{h^2}, -\frac{\nu}{h^2} + \frac{\beta}{2h} \right)$$



R. J. LeVeque, Finite Difference Methods for Ordinary and Partial Differential Equations, 2007, *SIAM*.

## ADVECTION DIFFUSION ODE – CROSS PRECONDITIONING

We consider the classical advection-diffusion ODE in one dimension, i.e.

$$\begin{aligned} -\nu \ddot{u} + \beta \dot{u} &= f \text{ in } (a, b) \subset \mathbb{R}, \\ u(a) &= 0, \quad u(b) = 1, \quad \nu, \beta \in \mathbb{R}_{\geq 0}. \end{aligned}$$

For the moment we will consider neither diffusion nor advection-dominated regimes, i.e.  $\nu \approx \beta$ , and discretisation over an equi-spaced mesh of step-size  $h$ . Such a discretisation results in the matrix

$$\underline{\underline{A}} = \text{tridiag} \left( -\frac{\nu}{h^2} - \frac{\beta}{2h}, \frac{2\nu}{h^2}, -\frac{\nu}{h^2} + \frac{\beta}{2h} \right)$$



R. J. LeVeque, Finite Difference Methods for Ordinary and Partial Differential Equations, 2007, *SIAM*.

$n$	QR	RQ	$Q(A^T A)^{1/2}$	$(AA^T)^{1/2} Q$
10	2	12	2	4
100	2	-	2	6
1000	2	-	2	7

**Table:** Comparison of the number of iterations for different preconditioners for the left preconditioned normal equation. The CGNE method was terminated when the absolute residual was less than  $10^{-12}$ . If the method did not converge in 1000 iterations, we marked the number of iterations with a dash.



# ADVECTION DIFFUSION PDE

We consider the classical advection-diffusion PDE in two dimensions, i.e.

$$\begin{aligned} \mathcal{L}u &:= -\nu \Delta u + \underline{\beta} \cdot \nabla u = f \text{ in } \Omega \subset \mathbb{R}^d, \\ u &= g \text{ on } \partial\Omega, \text{ with } \nu \ll \|\beta\|, \nabla \cdot \beta = 0. \end{aligned}$$



H. Elman, D. Silvester, A. Wathen,  
Finite Elements and Fast Iterative  
Solvers, 2005, *Oxford University  
Press*

# ADVECTION DIFFUSION PDE

We consider the classical advection-diffusion PDE in two dimensions, i.e.

$$\begin{aligned}\mathcal{L}u &:= -\nu \Delta u + \underline{\beta} \cdot \nabla u = f \text{ in } \Omega \subset \mathbb{R}^d, \\ u &= g \text{ on } \partial\Omega, \text{ with } \nu \ll \|\beta\|, \nabla \cdot \beta = 0.\end{aligned}$$



H. Elman, D. Silvester, A. Wathen,  
Finite Elements and Fast Iterative  
Solvers, 2005, *Oxford University  
Press*

## Finite Element Discretisation

Fix a discrete space  $V_h \subset H_0^1(\Omega)$  and look for  $u_h \in V_h$  such that

$$(\hat{\mathcal{L}}u_h, v_h) = \nu(\nabla u_h, \nabla v_h)_{L^2(\Omega)} + (\beta \cdot \nabla u_h, v_h)_{L^2(\Omega)} = (f, v_h)_{L^2(\Omega)} \text{ for any } v_h \in V_h.$$

# THE NORMAL EQUATIONS

---

We now need to understand what are the normal equations associated with the linear system,

$$A\underline{x} = \underline{b}, \text{ with } A_{ij} = (\hat{\mathcal{L}}\varphi_i, \varphi_j)_{L^2(\Omega)} \text{ and } b_j = (f, \varphi_j)_{L^2(\Omega)}.$$

# THE NORMAL EQUATIONS

We now need to understand what are the normal equations associated with the linear system,

$$A\underline{x} = \underline{b}, \text{ with } A_{ij} = (\hat{\mathcal{L}}\varphi_i, \varphi_j)_{L^2(\Omega)} \text{ and } b_j = (f, \varphi_j)_{L^2(\Omega)}.$$

The first thing we need to understand is what is  $\underline{\underline{A}}^T$ , in fact  $\underline{\underline{A}}^T$  is neither **Hilbert adjoint** of  $A$  nor the **Banach adjoint** seen as the operator  $A : V_h \subset H_0^1(\Omega) \rightarrow H^{-1}(\Omega) \subset V'_h$ .

## THE NORMAL EQUATIONS

We now need to understand what are the normal equations associated with the linear system,

$$A\underline{x} = \underline{b}, \text{ with } A_{ij} = (\hat{\mathcal{L}}\varphi_i, \varphi_j)_{L^2(\Omega)} \text{ and } b_j = (f, \varphi_j)_{L^2(\Omega)}.$$

The first thing we need to understand is what is  $\underline{\underline{A}}^T$ , in fact  $\underline{\underline{A}}^T$  is neither **Hilbert adjoint** of  $A$  nor the **Banach adjoint** seen as the operator  $A : V_h \subset H_0^1(\Omega) \rightarrow H^{-1}(\Omega) \subset V_h'$ .

In fact,  $A^T$  is an operator itself of the form  $A^T : V_h \subset H_0^1(\Omega) \rightarrow H^{-1}(\Omega) \subset V_h'$  which corresponds to the discretisation of the **Hilbert adjoint** of  $\mathcal{L}$ , i.e.

$$A_{ij}^T = A_{ji} = (\hat{\mathcal{L}}\varphi_j, \varphi_i)_{L^2(\Omega)} = (\varphi_j, \hat{\mathcal{L}}^*\varphi_i)_{L^2(\Omega)} = (\hat{\mathcal{L}}^*\varphi_i, \varphi_j)_{L^2(\Omega)},$$

## THE NORMAL EQUATIONS – PRIMAL DUAL ERROR

If we consider the classical normal equations, i.e.  $\underline{\underline{A}}^T \underline{\underline{A}} \underline{\underline{x}} = \underline{\underline{A}}^T \underline{\underline{b}}$ .

### Primal Dual Error

We notice that there is a primal dual error in the classical formulation of the normal equations.

$$V_h \subset H_0^1(\Omega) \xrightarrow{A} H^{-1} \subset V_h' \qquad V_h \subset H_0^1(\Omega) \xrightarrow{A^T} H^{-1} \subset V_h'$$

## THE NORMAL EQUATIONS – PRIMAL DUAL ERROR

If we consider the classical normal equations, i.e.  $\underline{\underline{A}}^T \underline{\underline{A}} \underline{\underline{x}} = \underline{\underline{A}}^T \underline{\underline{b}}$ .

### Primal Dual Error

We notice that there is a primal dual error in the classical formulation of the normal equations.

$$V_h \subset H_0^1(\Omega) \xrightarrow{A} H^{-1} \subset V_h' \qquad V_h \subset H_0^1(\Omega) \xrightarrow{A^T} H^{-1} \subset V_h'$$

To make sense of the normal equations we need to consider a Riesz map  $T : V_h' \rightarrow V_h$ .

$$V_h \subset H_0^1(\Omega) \xrightarrow{A} H^{-1} \subset V_h' \xrightarrow{T} V_h \subset H_0^1(\Omega) \xrightarrow{A^T} H^{-1} \subset V_h'$$

## THE NORMAL EQUATIONS

The Riesz map gives rise to a discrete operator  $T : V_h' \rightarrow V_h$ , which is **symmetric and positive definite**. Therefore if we consider the normal equations with respect to the Riesz map, i.e.

$$\underline{\underline{A}}^T T \underline{\underline{A}} \underline{x} = \underline{\underline{A}}^T T \underline{b},$$

we can rewrite them using a Cholesky factorisation of  $T$ , i.e.  $T = C^T C$ .

$$(CA)^T (CA) \underline{x} = (CA)^T C \underline{b},$$

hence the previous normal equations are associated with the linear system  $CA \underline{x} = C \underline{b}$ .

- ▶ The normal equations are still symmetric and positive definite. Hence we can solve them using CGNE. The cross-preconditioning idea is still applicable.
- ▶ The condition number of the normal equations is the square of the condition number of the original system.



## THE NORMAL EQUATIONS – $L^2$ -RIESZ MAP

---

We can consider as Riesz map the  $L^2$ -Riesz map, i.e.

$$(Tf, v_h)_{L^2(\Omega)} = \langle f, v_h \rangle \text{ for any } v_h \in V_h, f \in V'_h$$

## THE NORMAL EQUATIONS – $L^2$ -RIESZ MAP

We can consider as Riesz map the  $L^2$ -Riesz map, i.e.

$$(Tf, v_h)_{L^2(\Omega)} = \langle f, v_h \rangle \text{ for any } v_h \in V_h, f \in V_h'$$

Using the  $L^2$ -Riesz map the new normal is approximating, in the limit  $\nu \rightarrow 0$ , the problem: find  $u \in H_0^1(\Omega)$  such that

$$(\beta \otimes \beta \nabla u, \nabla v)_{L^2(\Omega)} = (g, v)_{L^2(\Omega)} \text{ for any } v \in H_0^1(\Omega).$$

## THE NORMAL EQUATIONS – $L^2$ -RIESZ MAP

We can consider as Riesz map the  $L^2$ -Riesz map, i.e.

$$(Tf, v_h)_{L^2(\Omega)} = \langle f, v_h \rangle \text{ for any } v_h \in V_h, f \in V'_h$$

Using the  $L^2$ -Riesz map the new normal is approximating, in the limit  $\nu \rightarrow 0$ , the problem: find  $u \in H_0^1(\Omega)$  such that

$$(\beta \otimes \beta \nabla u, \nabla v)_{L^2(\Omega)} = (g, v)_{L^2(\Omega)} \text{ for any } v \in H_0^1(\Omega).$$

$\nu$	CGNE Iterations
$1 \cdot 10^{-2}$	4231
$5 \cdot 10^{-3}$	3803
$2.5 \cdot 10^{-3}$	3327
$1.25 \cdot 10^{-3}$	2419

Table: The CGNE methods were terminated when the absolute residual was less than  $10^{-5}$ .

Due to the function space involved in the weak form, **we chose the wrong Riesz map.**

$$H_0^1(\Omega) \longrightarrow H^{-1} \subset L^{2'} \xrightarrow{T^{-1}} L^2 \not\subset H_0^1(\Omega) \longrightarrow H^{-1}$$

## THE NORMAL EQUATIONS – $H^1$ -RIESZ MAP

---

We can consider as Riesz map the  $H^1$ -Riesz map, i.e.

$$(\nabla T f, \nabla v_h)_{L^2(\Omega)} = \nu^{-1} \langle f, v_h \rangle, \quad \forall v_h \in V_h, f \in V_h'.$$

## THE NORMAL EQUATIONS – $H^1$ -RIESZ MAP

We can consider as Riesz map the  $H^1$ -Riesz map, i.e.

$$(\nabla T f, \nabla v_h)_{L^2(\Omega)} = \nu^{-1} \langle f, v_h \rangle, \quad \forall v_h \in V_h, f \in V_h'.$$

Using this Riesz map the normal equations  $\underline{\underline{A}}^T T \underline{\underline{A}} \underline{x} = \underline{\underline{A}}^T T \underline{b}$  is approximating the problem: find  $u \in H_0^1(\Omega)$  such that

$$\nu(\nabla u, \nabla v)_{L^2(\Omega)} + \nu^{-1}(\Pi_{\nabla} \beta u, \Pi_{\nabla} \beta v)_{L^2(\Omega)}, \quad \text{for any } v \in H_0^1(\Omega).$$

## THE NORMAL EQUATIONS – $H^1$ -RIESZ MAP

We can consider as Riesz map the  $H^1$ -Riesz map, i.e.

$$(\nabla T f, \nabla v_h)_{L^2(\Omega)} = \nu^{-1} \langle f, v_h \rangle, \quad \forall v_h \in V_h, f \in V_h'.$$

$\nu$	$32 \times 32$	$64 \times 64$	$128 \times 128$
$1 \cdot 10^{-2}$	2	2	2
$5 \cdot 10^{-3}$	3	3	3
$2.5 \cdot 10^{-3}$	3	3	3
$1.25 \cdot 10^{-3}$	3	3	3

**Table:** The CGNE methods were terminated when the absolute residual was less than  $10^{-5}$ .

Using this Riesz map the normal equations  $\underline{\underline{A}}^T T \underline{\underline{A}} \underline{x} = \underline{\underline{A}}^T T \underline{b}$  is approximating the problem: find  $u \in H_0^1(\Omega)$  such that

$$\nu(\nabla u, \nabla v)_{L^2(\Omega)} + \nu^{-1}(\Pi_{\nabla} \beta u, \Pi_{\nabla} \beta v)_{L^2(\Omega)}, \quad \text{for any } v \in H_0^1(\Omega).$$

# THE NORMAL EQUATIONS – PRECONDITION USING THE MASS MATRIX AND AMG

Find  $u_h \in V_h$  such that  $\nu^{-1}(\beta u_h, \beta v_h)_{L^2(\Omega)}$ , for any  $v_h \in V_h$ .

$\nu$	$32 \times 32$	$64 \times 64$	$128 \times 128$	$256 \times 256$
$1 \cdot 10^{-2}$	10	15	20	23
$5 \cdot 10^{-3}$	11	15	22	30
$2.5 \cdot 10^{-3}$	17	16	21	32
$1.25 \cdot 10^{-3}$	26	24	23	30

Table: Comparison of the number of iterations for the CGNE method preconditioned by the inversion via PETSc GAMG, for different values of  $\nu$  and different mesh sizes. The wind is fixed to  $\sqrt{2}\beta = (1, 1)$  and as right-hand side we consider the function  $f(x, y) \equiv 1$ . The CGNE method was terminated when the absolute residual was less than  $10^{-5}$ .

# THE NORMAL EQUATIONS – PROJECTED MASS MATRIX AND GMG

Find  $u_h \in V_h$  such that  $\nu^{-1}(\Pi_{\nabla}\beta u_h, \Pi_{\nabla}\beta v_h)_{L^2(\Omega)}$ , for any  $v_h \in V_h$ .

$\nu$	$32 \times 32$	$64 \times 64$	$128 \times 128$
$1 \cdot 10^{-2}$	4	5	8
$5 \cdot 10^{-3}$	4	5	7
$2.5 \cdot 10^{-3}$	5	5	7
$1.25 \cdot 10^{-3}$	7	7	7

Table: Comparison of the number of iterations for the CGNE method preconditioned by geometric multigrid with SOR smoothing, for different values of  $\nu$  and different mesh sizes. The wind is fixed to  $\sqrt{2}\beta = (1, 1)$  and as right-hand side we consider the function  $f(x, y) \equiv 1$ . The CGNE method was terminated when the absolute residual was less than  $10^{-5}$ .



## TAKE AWAY MESSAGE

---

- ▶ The normal equations are a powerful tool to solve linear systems arising from PDEs, for which we have a very good understanding of convergence.

## TAKE AWAY MESSAGE

---

- ▶ The normal equations are a powerful tool to solve linear systems arising from PDEs, for which we have a very good understanding of convergence.
- ▶ The correct notion of a good preconditioner for the normal equations is crucial to understand how to precondition the normal equations. We propose the notion of **cross preconditioning**.

## TAKE AWAY MESSAGE

---

- ▶ The normal equations are a powerful tool to solve linear systems arising from PDEs, for which we have a very good understanding of convergence.
- ▶ The correct notion of a good preconditioner for the normal equations is crucial to understand how to precondition the normal equations. We propose the notion of **cross preconditioning**.
- ▶ A careful study of the normal equations can suggest a new PDE to use as preconditioner. Often these PDEs are simpler to solve than the original ones. We refer to this idea as **normal preconditioning**.

## TAKE AWAY MESSAGE

---

- ▶ The normal equations are a powerful tool to solve linear systems arising from PDEs, for which we have a very good understanding of convergence.
- ▶ The correct notion of a good preconditioner for the normal equations is crucial to understand how to precondition the normal equations. We propose the notion of **cross preconditioning**.
- ▶ A careful study of the normal equations can suggest a new PDE to use as preconditioner. Often these PDEs are simpler to solve than the original ones. We refer to this idea as **normal preconditioning**.
- ▶ **We should reconsider the use of normal equations for solving linear systems arising from PDEs.**

## FUTURE WORK

---

- ▶ There is an intimate connection between the notion of **normal preconditioning** and a method known as **discontinuous Petrov-Galerkin**. We would like to further explore this connection and understand the optimisation problem associated with the normal equations here proposed.

## FUTURE WORK

---

- ▶ There is an intimate connection between the notion of **normal preconditioning** and a method known as **discontinuous Petrov-Galerkin**. We would like to further explore this connection and understand the optimisation problem associated with the normal equations here proposed.
- ▶ Explore the notion of **normal preconditioning** for higher-order finite element discretisation.

## FUTURE WORK

---

- ▶ There is an intimate connection between the notion of **normal preconditioning** and a method known as **discontinuous Petrov-Galerkin**. We would like to further explore this connection and understand the optimisation problem associated with the normal equations here proposed.
- ▶ Explore the notion of **normal preconditioning** for higher-order finite element discretisation.
- ▶ Apply **normal preconditioning** to other PDEs such as the Helmholtz equation, using as Riesz map the T-coercive map. We would also like to study the Oseen equation and  $C^1$  nearly singular problems such as the Helmholtz–Korteweg equation.

**THANK YOU!**  
**Lorenzo now accepts questions.**

---

Some more comments on the normal equations: With a focus on discretisation of partial differential equations

L. LAZZARINO, Y. NAKATSUKASA, UMBERTO ZERBINATI\*