ILLINOIS

Bioengineering Department
GRAINGER COLLEGE OF ENGINEERING

Learning a Projection Operator onto the Null Space of a Linear Imaging Operator

Albert Zhai

BIOE 580

November 29, 2021

Outline

- Imaging operators and null spaces
- What is P_{null} and why is it useful?
- Difficulties in computing P_{null}
- Optimization-based methods for obtaining P_{null}
- Experimental results
- Next steps

Imaging operators and null spaces

$$g = \mathcal{H} f$$

- A digital imaging system can be modeled as a linear continuous-to-discrete (CD) operator
- In practice, approximated as discrete-to-discrete (DD):

$$\mathcal{H} \in \mathbb{E}^{M \times N}$$

- Null space:

$$N(\mathcal{H}) = \{f \in \mathbb{E}^N | \mathcal{H} f = 0\}$$

What is P_{null} and why is it useful?

- Objects in null space are “invisible” to imaging system
  - Measurable space := $N_{\perp}(\mathcal{H})$
- Any object can be uniquely expressed as the sum of its **measurable component** and its **null component**:

$$f = f_{meas} + f_{null}$$

$$f_{meas} \in N_{\perp}(\mathcal{H}), \quad f_{null} \in N(\mathcal{H})$$

What is P_{null} and why is it useful?

- P_{null} is an orthogonal projection onto the null space
- Can be used to extract null component of any object:

$$P_{null} f = f_{null}$$

- Analysis of null components allows us to assess what objects features are invisible to an imaging system
  - Can optimize system to reduce amount of task-relevant features that are missed
- Can also assess hallucinations in reconstructions

Difficulties in computing P_{null}

- Dimensionality of $\mathcal{H} \in \mathbb{E}^{M \times N}$ often very large
  - E.g. M ≈ N = 1024x1024 → H has 10^{12} entries
- Explicit storage and matrix-multiplication are prohibitively expensive
- Classical linear algebra methods become infeasible:

$$P_{null} = I - \mathcal{H}^+ \mathcal{H}$$

$$\mathcal{H}^+ = U \Sigma^+ V^*$$

Optimization-based methods for obtaining P_{null}

- For any f,

$$P_{null} f = f_{null} \qquad \mathcal{H} f_{null} = 0$$

- Let R be the rank of H. If we have

$$W \in \mathbb{E}^{N \times (N-R)} \;\; s.t. \qquad W^* W = I$$

and

$$W = \operatorname{argmin}_W ||\mathcal{H} W W^* f||_2^2$$

for all f, then

$$W W^* = P_{null}$$

Optimization-based methods for obtaining P_{null}

- How to ensure

$$W = \text{argmin}_W ||\mathcal{H}WW^* f||_2^2$$

  for all f?

- Treat f as a random variable and minimize the expectation:

$$W = \text{argmin}_W \mathbb{E}_f [||\mathcal{H}WW^* f||_2^2]$$

  - Support of distribution must span object space
  - Draw samples and perform SGD

Optimization-based methods for obtaining P_{null}

$$W \in \mathbb{E}^{N\times(N-R)} \;\; s.t. \qquad P_{null} = WW^*$$

Optimization-based methods for obtaining P_{null}

$$\cancel{W \in \mathbb{E}^{N\times(N-R)} \quad s.t. \quad P_{null} = WW^*}$$

$$W \in \mathbb{E}^{M\times R} \quad s.t. \quad \begin{aligned} P_{null} &= I - P_{meas} \\ P_{meas} &= \mathcal{H}^* WW^* \mathcal{H} \end{aligned}$$

- M < N
- Range of projection is already restricted to $N_{\perp}(\mathcal{H})$

Optimization-based methods for obtaining P_{null}

- If $\mathcal{H}^* W$ is orthogonal, then $\mathcal{H}^* W W^* \mathcal{H}$ has rank R
- This guarantees that $\mathcal{H}^* W W^* \mathcal{H} = P_{meas}$
- All we need is

$$W^* \mathcal{H}\mathcal{H}^* W = I$$

- Solve minimization problem:

$$W = \mathrm{argmin}_W ||W^* \mathcal{H}\mathcal{H}^* W - I||_2^2$$

- To avoid matrix-multiplication, minimize indirectly:

$$W = \mathrm{argmin}_W \mathbb{E}_v [||(W^* \mathcal{H}\mathcal{H}^* W - I)v||_2^2]$$

Experiments on RT operator

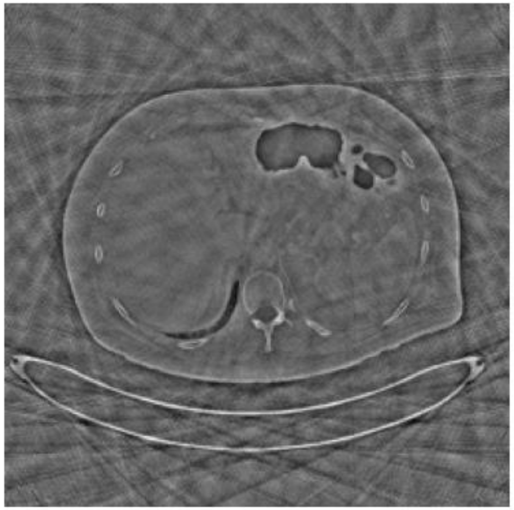
- 2D fan-beam curved-detector tomography
  - ***N* = 256x256 = 65536**
  - ***M* = 10860**
    - 30 views
    - round(256 sqrt(2)) = 362 rays per view
  - ***R* = 10362**
- Train on white noise: $v \sim \mathcal{N}(0, I)$
  - Batch size 512
  - Adam optimizer with decaying learning rate

Experimental results

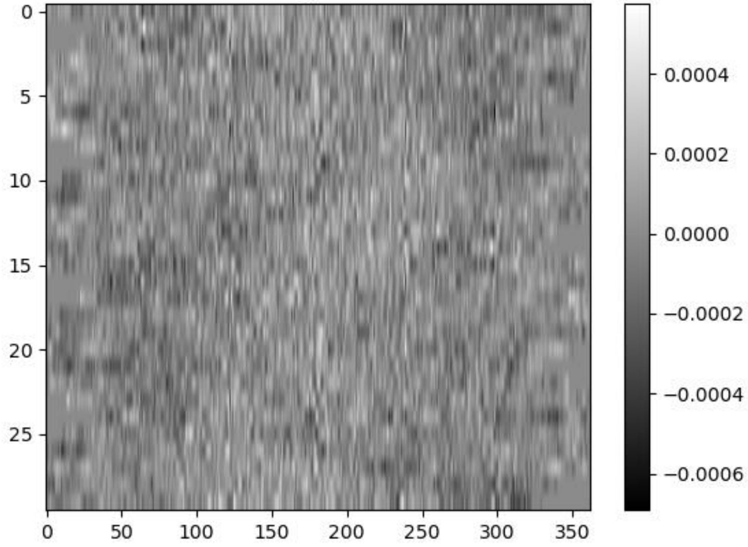
- $I - \mathcal{H}^* W W^* \mathcal{H}$ method is significantly more efficient but cannot converge to very low error

	Peak memory
rSVD	34.4 GB
WW^*	21.3 GB
$I - \mathcal{H}^* W W^* \mathcal{H}$	4.25 GB

f_{null}



$\mathcal{H} f_{null}$



Next steps

- Want to achieve better convergence
- Investigate different training distributions
  - Sample from fixed orthogonal basis
  - Going back to $W = \text{argmin}_W \mathbb{E}_f[||\mathcal{H}WW^* f||_2^2]$ and sampling real medical data