Locally Sparse Reconstruction Using $\ell^{1,\infty}$-Norms

Inverse Days 2013, Inari, Finland
Introduction
Motivation
Inverse Problem

Variational Model

Some Computational Results

Summary & Outlook
Introduction

Motivation

Inverse Problem

Variational Model

Some Computational Results

Summary & Outlook
Myocardial Infraction

- Many people suffer from coronary heart disease
- Heart attacks are life-threatening and often lead to death
- Necessity to know how well the cardiac muscle is perfused and where exactly the problems are located

Heart With Muscle Damage and a Blocked Artery

Figure: U.S. National Heart Lung and Blood Institute

Pia Heins (pia.heins@wwu.de), Martin Burger
Visualize the Myocardial Perfusion

- Visualize the **perfusion** of the cardiac muscle
- Locate **damaged areas** after a heart attack
- Find arteries which might become a problem to **prevent an infarction**

**Figure:** U.S. National Heart Lung and Blood Institute
Use Dynamic Positron-Emission-Tomography

- PET is an **imaging technique** in nuclear medicine
- It visualizes the **distribution** of a weak radioactive labeled substance (**tracer**) in order to image functional processes in the body

**Figure**: J. Langner, M.Sc. Thesis, 2003
Inverse Problem

From dynamic PET we obtain the following discrete \textit{time-dependent} inverse problem

\[ AZ = W \]

where

- \( A \in \mathbb{R}^{L \times M} \) is the \textbf{PET-Matrix}
- \( Z \in \mathbb{R}^{M \times T} \) is the \textbf{unknown dynamic image}, i.e. the distribution of the tracer over the time
- \( W \in \mathbb{R}^{L \times T} \) is the \textbf{measured data}

with \( M \) pixel, \( T \) time steps and \( L \) bins.
Myocardial Perfusion with Dynamic PET

- **Tracer:** $H_{2}^{15}O$ (short half-life $\sim$ bad statistics)
- **Data:** partitioned in temporal bins
- **First approach:** generate sequence of images from the temporal bins

**Figure:** M. Benning, Improved Dynamic PET via Kinetic Models and Operator Splitting, 2010
Problem

Due to the **bad statistics**, we obtain **bad quality images**.

We also **lose temporal information**.

**Figure:** Transversal Cut
Include Kinetic Modeling\(^1\)

- Model the **tracer exchange** in the capillaries
- Obtain equation for the **dynamic image**, which depends on some parameters
- **Parameters** are in a **certain range**

\(^1\)[3, Wernick & Aarsvold, Emission Tomography, 2004]
Kinetic Modeling Basis Functions

- Choose different parameters out of the reasonable interval
- Obtain basis vectors from different parameters in the blood flow model
- Basis vectors $b_n$ are computable in advance
Inverse Problem including Dictionary

Assume that every pixel $m$ at time step $t$ of the image $Z$ can be written as a linear combination of known basis vectors $b_t$ with coefficient vectors $u_m$, i.e.

$$z_{mt} = \sum_{n=1}^{N} u_{mn} b_{tn} \quad \Rightarrow \quad Z = UB^T.$$

Thus we obtain

$$AUB^T = W$$

with $A \in \mathbb{R}^{L \times M}$, $U \in \mathbb{R}^{M \times N}$, $B \in \mathbb{R}^{T \times N}$ and $W \in \mathbb{R}^{L \times T}$.

There are other applications, which lead to a similar problem, i.e. unmixing problems.
Knowledge about the Basis

- **every pixel** should consist of **only one** or at most **very few** of the given basis vectors

- consider the matrix $B$ to be **coherent**, i.e. the coherence parameter
  $$
  \mu := \max_{i \neq j} |\langle b_i, b_j \rangle|
  $$

  for $b_i, b_j$ being distant basis vectors, is large.

  In other words, the **basis vectors are very similar**.

Note that orthogonalization does not help because the coefficients would not be sparse and we would lose our prior knowledge.
Introduction
Motivation
Inverse Problem

Variational Model

Some Computational Results

Summary & Outlook
A Priori Knowledge

Variational Problem

\[ \frac{1}{2} \| AUB^T - W \|_F^2 + \alpha \mathcal{R}(U) \rightarrow \min_U \]

Prior Knowledge:
(At best) we would like to have just one (or only a few) basis vector, which fits best in the considered pixel.

\[ \Rightarrow \] only one coefficient unequal to zero per pixel.

\[ \Rightarrow \] want to promote sparsity in every pixel.
Which Regularization?

Idea: $\ell^{0,\infty}$-Regularization

$$\min_U \|U\|_{0,\infty} = \min_U \left\{ \max_i \sum_{j=1}^N |u_{ij}|^0 \right\}$$
Which Regularization?

Idea: $\ell_1, \infty$-Regularization as Relaxation

$$\min_U \| U \|_{1, \infty} = \min_U \left\{ \max_i \sum_{j=1}^{N} |u_{ij}| \right\}$$

$^{3}$[1, Donoho & Elad, Optimally Sparse Representation in General (non-Orthogonal) Dictionaries via l1 Minimization, 2003]
Which Regularization?

Idea: $\ell^{1,\infty}$-Regularization as Relaxation

\[
\min_U \|U\|_{1,\infty} = \min_U \left\{ \max_i \sum_{j=1}^N |u_{ij}| \right\}
\]

Variational Model

\[
\min_U \frac{1}{2} \|AUB^T - W\|_F^2 + \alpha \|U\|_{\ell^{1,\infty}}
\]

[1, Donoho & Elad, Optimally Sparse Representation in General (non-Orthogonal) Dictionaries via l1 Minimization, 2003]
Reconstruction with Local Sparsity

Implementation

- The implementation of the $\ell^{1,\infty}$-term and also its analysis is not that easy

- Reformulation of the problem is necessary

- Simplified assumption: Nonnegativity of the coefficients (reasonable in many applications)
Reformulation

\[
\min_U \frac{1}{2} \|AUB^T - W\|_F^2 + \alpha \max_i \sum_{j=1}^{N} |u_{ij}|
\]
Reformulation

Add a nonnegativity constraint

\[
\min_U \frac{1}{2} \|AUB^T - W\|_F^2 + \alpha \max_i \sum_{j=1}^{N} u_{ij} \quad \text{s. t.} \quad u_{ij} \geq 0 \quad \forall \ i, j
\]
Reformulation

Add a nonnegativity constraint

$$\min_u \frac{1}{2} \| AUB^T - W \|_F^2 + \alpha \max_i \sum_{j=1}^{N} u_{ij} \quad \text{s. t.} \quad u_{ij} \geq 0 \quad \forall i, j$$

Maximum is a problem!
Reformulation

Add a nonnegativity constraint

\[
\min_{U} \quad \frac{1}{2} \|AUB^T - W\|_F^2 + \alpha \max_{i} \sum_{j=1}^{N} u_{ij} \quad \text{s. t.} \quad u_{ij} \geq 0 \quad \forall i, j
\]

Maximum is a problem!

\[\Rightarrow\] Use equivalent formulation, i.e.

\[
\min_{U, \tilde{v}} \quad \frac{1}{2} \|AUB^T - W\|_F^2 + \tilde{v} \quad \text{s. t.} \quad \alpha \sum_{j=1}^{N} u_{ij} \leq \tilde{v}, \quad u_{ij} \geq 0 \quad \forall i, j
\]
Reformulated Problems

Analog to the unstrained problem

\[
\begin{align*}
\min_{U, \tilde{v}} & \quad \frac{1}{2} \| AUB^T - W \|_F^2 + \tilde{v} \\
\text{s. t.} \quad & \quad \alpha \sum_{j=1}^{N} u_{ij} \leq \tilde{v}, \quad u_{ij} \geq 0 \quad \forall \ i, j
\end{align*}
\]

we obtain the constrained problem

\[
\begin{align*}
\min_{U, \tilde{v}} & \quad \tilde{v} \\
\text{s. t.} \quad & \quad \alpha \sum_{j=1}^{N} u_{ij} \leq \tilde{v}, \quad u_{ij} \geq 0 \quad \forall \ i, j, \ AUB^T = W
\end{align*}
\]

on which we did some analysis about exact recovery.
Locally 1-Sparse Exact Recovery

**Sufficient Condition** (*Constrained Problem*)

In case

1. the exact solution has only **one non-zero entry** per pixel,

2. the matrix $B$ is normalized with a **strict inequality** for distinct basis vectors,

3. we **include** an additional $\ell^{1,1}$-regularization,

then we are able to **reconstruct the support exactly**, i.e. we only have to deal with loss of contrast.
Further Modification of the Problem

\[
\min_{U, \tilde{v}} \frac{1}{2} \|AU^T - W\|_F^2 + \tilde{v} \quad \text{s. t.} \quad \alpha \sum_{j=1}^{N} u_{ij} \leq \tilde{v}, \quad u_{ij} \geq 0 \quad \forall \ i, j
\]
Further Modification of the Problem

Add $\ell^{1,1}$-regularization.

$$\min_{U, \tilde{v}} \frac{1}{2} \| AUB^T - W \|_F^2 + \tilde{v} + \beta \| U \|_{\ell^{1,1}} \quad \text{s. t.} \quad \alpha \sum_{j=1}^{N} u_{ij} \leq \tilde{v}, \quad u_{ij} \geq 0 \quad \forall \ i, j$$
Further Modification of the Problem

Add $\ell^{1,1}$-regularization.

The problem is still difficult to implement.

$$\min_{U, \tilde{v}} \frac{1}{2} \|AUB^T - W\|_F^2 + \tilde{v} + \beta \|U\|_{\ell^{1,1}} \quad \text{s. t.} \quad \sum_{j=1}^{N} u_{ij} \leq \frac{\tilde{v}}{\alpha}, \quad u_{ij} \geq 0 \quad \forall \ i, j$$
Further Modification of the Problem

Add $\ell^{1,1}$-regularization.

The problem is still difficult to implement.

$$\min_{U} \frac{1}{2} \|AUB^T - W\|_F^2 + \beta \|U\|_{\ell^{1,1}} \quad \text{s. t.} \quad \sum_{j=1}^{N} u_{ij} \leq v, \quad u_{ij} \geq 0 \quad \forall i, j$$

Instead of regularizing with $\alpha$ and minimizing over $\tilde{v}$, we can choose $v$ in advance and thus *regularize with $v$* instead.

Note that we make a **systematic error** and only **obtain the support**.
Introduction
Motivation
Inverse Problem

Variational Model

Some Computational Results

Summary & Outlook
Algorithm

Splitting Approach:

\[ \min_{U, Z, D} \frac{1}{2} \|AZ - W\|_F^2 + \beta \sum_{i,j} d_{ij} \quad \text{s. t.} \quad \sum_{j=1}^{N} d_{ij} \leq v, \quad d_{ij} \geq 0 \quad \forall i \]

\[ Z = UBT, \quad D = U \]

Solve via Alternating Direction Method of Multipliers\(^3\) (ADMM), i.e.

- state Augmented Lagrangian
- compute optimality conditions
- solve subproblems successively

\(^3\)[2, D. Gabay, Applications of the Method of Multipliers to Variational Inequalities]
Exact Coefficients

- Use simple 3D matrix $\hat{U}$ containing the exact coefficients
- Define 2 regions where coefficients are nonzero for one basis vector
- Thus the corresponding coefficients for the most basis vectors are zero
Basic Idea

Strong regularization $\implies$ very good reconstruction of the support

Figure: Reconstruction with $\nu = 0.1$ and $\beta = 0.1$

Every value larger than $\nu$ is projected down to $\nu$
$\implies$ we are not really close to the exact data
Basic Idea

- **First run** with $\ell^{1,\infty}$- and $\ell^{1,1}$-regularization to obtain the support
Basic Idea

- **First run** with $\ell^{1,\infty}$- and $\ell^{1,1}$-regularization to obtain the support
- **Second run** without regularization *only on the known support* to reduce the distance to the exact data

---

Pia Heins (pia.heins@wwu.de), Martin Burger
Basic Idea

▶ **First run** with $\ell_1,\infty$ - and $\ell_1,1$-regularization to obtain the support

▶ **Second run** without regularization *only on the known support* to reduce the distance to the exact data

**Figure:** Reconstruction with $\nu = 0.1$ and $\beta = 0.1$ including second run

⇝ very good results
Example including Noise

Figure: Reconstruction of a $200 \times 200 \times 8$ image with $\nu = 0.01$ and $\beta = 0.1$ and standard deviation $\sigma = 0.01$
Example including Noise

More regularization
⇝ prior knowledge is fulfilled
Example including Noise

More regularization
\( \Rightarrow \) prior knowledge is fulfilled

Basis functions are very similar
\( \Rightarrow \) in some pixels we obtain the “wrong” basis function

Reminder:

Input Curve \( C_A(t) \) and Kinetic Modeling Basis Functions \( b_n(t) \)
How to make further improvements?

Use **additional** problem-specific **regularization in space**

~~> include **total variation** on the image

\[
\min_U \frac{1}{2} \| AUB^T - W \|_F^2 + \alpha \| U \|_{\ell^1,\infty} + \beta \| U \|_{\ell^1,1} + \gamma \| UB^T \|_{TV}
\]

~~> still **work in progress**
Summary

- Similar problem in different applications (dPET, FLIM, ECG, ...)
  - use specific operator and basis functions

- Reformulated problem is easier to implement and leads to the same solution

- Use ADMM for the double splitting

- Exact recovery of the support under certain circumstances
Outlook

- Get TV-regularization to work properly
- Use more difficult data, i.e. more and smaller regions
- Use larger data
Thank you for your attention!

Questions?
Bibliography

D. L. Donoho and M. Elad.
Optimally sparse representation in general (non-orthogonal) dictionaries via $l_1$ minimization.


D. Gabay.
Applications of the method of multipliers to variational inequalities.


M. N. Wernick and J. N. Aarsvold, editors.
*Emission Tomography: The Fundamentals of PET and SPECT.*
Sufficient Condition for Locally 1-Sparse Exact Recovery

Consider the constrained problem, i.e.

$$\min_{U, \tilde{v}} \quad \beta \|U\|_{1,1} + \tilde{v} \quad \text{s.t.} \quad \alpha \sum_{j=1}^{N} u_{ij} \leq \tilde{v}, \quad u_{ij} \geq 0 \quad \forall \ i, j, \ AUB^T = W. \quad (1)$$

For this problem the Lagrange functional reads as follows:

$$\mathcal{L}(v, u_{ij}; \lambda, \mu, \eta) = \beta \sum_{i=1}^{M} \sum_{j=1}^{N} u_{ij} + v + \sum_{i=1}^{M} \lambda_i \left( \alpha \sum_{j=1}^{N} u_{ij} - v \right)$$

$$- \sum_{j=1}^{N} \sum_{i=1}^{M} \mu_{ij} u_{ij} + \sum_{l=1}^{L} \sum_{k=1}^{T} \eta_{lk} \left( w_{lk} - \sum_{j=1}^{N} \sum_{i=1}^{M} a_{li} u_{ij} b_{kj} \right),$$

where $\lambda$, $\mu$, and $\eta$ are Lagrange parameters.
Sufficient Condition for Locally 1-Sparse Exact Recovery

Want to use the \textit{strong} scaling condition

\[
\left\| \sqrt{\delta} b_{J(i)} \right\|_2 = 1 \quad \text{and} \quad \left| \left\langle \sqrt{\delta} b_{J(i)}, \sqrt{\delta} b_j \right\rangle \right| < 1 \quad \forall \ i, j .
\]  

(2)

Set the exact solution of (1) as

\[
\hat{u}_{ij} = \begin{cases} 
    c_i, & \text{if } j = J(i) \\
    0, & \text{if } j \neq J(i) 
\end{cases} .
\]

Choose the Lagrange parameter \( \lambda \) as

\[
\lambda_i = \begin{cases} 
    \frac{1}{m}, & \text{if } c_i = \frac{v}{\alpha} \\
    0, & \text{if } c_i < \frac{v}{\alpha} 
\end{cases} \quad \forall \ i ,
\]

with \( \frac{v}{\alpha} = \max_p c_p \) and \( m \) being the number of indices for which \( c_i = \frac{v}{\alpha} \) holds.
Sufficient Condition for Locally 1-Sparse Exact Recovery

Choose the Lagrange parameter $\eta$ as

$$\eta_{lk} = \delta_k \sum_{n=1}^{N} (\Phi_{ln} + \beta) b_{kn} \quad \text{with} \quad A^T \Phi = \alpha \text{diag}(\lambda) \Psi,$$

with $\beta > 0$, $\delta_k \in \mathbb{R}$ and $\Psi$ defined via the indicator function of $\hat{u}_{ij}$:

$$\Psi_{ij} := \chi_{\hat{u}_{ij}} = \begin{cases} 1 & \text{if } j = J(i) \\ 0 & \text{if } j \neq J(i) \end{cases} \quad \forall \ i, j.$$

Let the strong scaling condition (2) hold. Then

$$u_{ij} = \begin{cases} \tilde{c}_i = c_i + r_i, & \text{if } j = J(i) \\ 0, & \text{if } j \neq J(i) \end{cases}$$

is a solution of (1) with $r_i \in \mathbb{R}$. 

Pia Heins (pia.heins@wwu.de), Martin Burger
Sufficient Condition for Locally 1-Sparse Exact Recovery

Sketch of Proof

- Set up the Lagrange-Functional
- Compute optimality and complementary conditions
- Satisfy optimality conditions
  - case-by-case analysis for \( j = J(i) \) and \( j \neq J(i) \)
  - insert Lagrange parameters
  - insert scaling condition to compute the last Lagrange parameter \( \mu_{ij} \)
  - use complementary condition to obtain the reconstruction \( u \)
Sufficient Condition for Locally $1$-Sparse Exact Recovery

**Sketch of Proof**

- Set up the Lagrange-Functional
- Compute optimality and complementary conditions
- Satisfy optimality conditions
  - case-by-case analysis for $j = J(i)$ and $j \neq J(i)$
  - insert Lagrange parameters
  - insert scaling condition to compute the last Lagrange parameter $\mu_{ij}$
  - use complementary condition to obtain the reconstruction $u$

- $\ell^{1,1}$-regularization to exclude dividing by zero
- strict inequality to deduce that $\mu_{ij} \neq 0$ for one Lagrange parameter. Thus $u_{ij} = 0$ holds for all $j \neq J(i)$