Tomography and Regularization

Luca Ratti
Department of Mathematics and Statistics, University of Helsinki
luca.ratti@helsinki.fi

Hiroshima, January 15, 2020
Review

Linear inverse problem

Determine $f$ given $y^\delta = Kf + \epsilon; \|\epsilon\| \leq \delta$

Stability (which usually fails)

Small perturbations on the data ($\delta$) should imply small perturbations on the solution ($f$).

Singular Value Decomposition

$K = U\Sigma V^T$, being $\Sigma$ a diagonal matrix and $U, V$ orthogonal matrices. Define the pseudoinverse $K^\dagger = V\Sigma^{-1}U^T$.

Truncated SVD

Instead of $K^\dagger y^\delta = \sum_{i=1}^{\min(m,n)} \frac{u_i^T y^\delta}{\sigma_i} v$, use $f^{\text{TSVD}} = \sum_{i=1}^{r} \frac{u_i^T y^\delta}{\sigma_i} v$, $r < \min(m,n)$. 
TSVD as Spectral Filtering

We can regard the TSVD also as the result of a filtering operation, namely:

\[ f^{\text{TSVD}} = \sum_{i=1}^{r} \frac{u_i^T y^\delta}{\sigma_i} v_i = \sum_{i=1}^{\min(m,n)} \phi^{\text{TSVD}}(\sigma_i) u_i^T y^\delta v_i \]

where \( r \) is the truncation parameter and

\[ \phi^{\text{TSVD}}(\sigma_i) = \begin{cases} \frac{1}{\sigma_i} & i = 1, \ldots, r \\ 0 & \text{elsewhere} \end{cases} \]

is the filter function associated with the method.

This is called spectral filtering methods because the SVD basis can be regarded as a spectral basis, since the vectors \( u_i \) and \( v_i \) are the eigenvectors of \( K^T K \) and \( K K^T \).
The Tikhonov Method

Let’s now consider the following filter function:

\[ \phi_{\text{Tikh}}(\sigma_i) = \frac{\sigma_i}{\sigma_i^2 + \alpha^2}, \quad i = 1, \ldots, \min(m,n) \]

which yield the reconstruction method:

\[ f_{\text{Tikh}} = \sum_{i=1}^{\min(m,n)} \phi_{\text{Tikh}}(\sigma_i) u_i^T y \delta v_i = \sum_{i=1}^{\min(m,n)} \frac{\sigma_i (u_i^T y \delta)}{\sigma_i^2 + \alpha^2} v_i. \]

This choice of the filter results in a regularization technique called **Tikhonov method** and \( \alpha > 0 \) is the so-called **regularization parameter**.

Comparison with **TSVD**: 

![Graph comparing Tikhonov and TSVD methods](image-url)
Tikhonov Regularization

It is possible to prove that \( f^{Tikh} \) can be obtained as the solution of the minimization problem:

\[
f^{Tikh} = \arg\min_{f} \left\{ \|Kf - y^\delta\|^2_2 + \alpha \|f\|^2_2 \right\}.
\]

Note that by selecting \( \alpha = 0 \) we would get the Moore-Pensore solution \( f^\dagger \)

\[
\|f^\dagger\|^2_2 = \sum_{i=1}^{k} \frac{(u_i^T y^\delta)^2}{\sigma_i^2},
\]

which we know to be unstable whenever the magnitude of the noise in some direction \( u_i \) greatly exceeds the magnitude of the singular value \( \sigma_i \).

By taking \( \alpha \to \infty \), instead, the solution of the minimization problem tends to \( f = 0 \): Tikhonov regularization penalizes solutions with large norms/ promotes solutions with small norms.

The regularization parameter balances the trade-off between the two building blocks of Tikhonov regularization: the data fidelity (\( \|Kf - y^\delta\| \)) and the a priori knowledge of the solution (in particular, its norm \( \|f\| \) should be small).
Influence of the Choice of $\alpha$ in Tikhonov Regularization

Original phantom

$f^\text{Tikh}: \alpha = 10^3$
Influence of the Choice of $\alpha$ in Tikhonov Regularization

Original phantom

$f^\text{Tikh}: \quad \alpha = 10^2$
Influence of the Choice of $\alpha$ in Tikhonov Regularization

Original phantom

$f^{\text{Tikh}}: \alpha = 10$
Influence of the Choice of $\alpha$ in Tikhonov Regularization

Original phantom

$f^{\text{Tikh}}: \alpha = 1$
Influence of the Choice of $\alpha$ in Tikhonov Regularization

Original phantom

$f^{\text{Tikh}}: \alpha = 10^{-1}$
Influence of the Choice of \( \alpha \) in Tikhonov Regularization

Original phantom

\( f^{\text{Tikh}} : \alpha = 10^{-2} \)
Influence of the Choice of $\alpha$ in Tikhonov Regularization

Original phantom

$f^\text{Tikh}: \alpha = 10^{-3}$
About the Regularization Parameter

Let’s summarize the effects of the choice of the parameter $\alpha$ on the solution of Tikhonov regularization:

$$f^{\text{Tikh}} = \arg\min_f \left\{ \|Kf - y^\delta\|_2^2 + \alpha \|f\|_2^2 \right\}.$$  

- A large $\alpha$ results in strong regularity and possible over smoothing
- A small $\alpha$ small yields a good fitting, with the risk of over fitting.

In general, choosing the regularization parameter for an ill-posed problem is not a trivial task and there are no rule of thumbs. Usually, it is a combination of good heuristics and prior knowledge of the noise in the observations.

Delving into this is out of the scope, but there are methods that can be found in the literature (Morozov’s discrepancy principle, generalized cross validation, L-curve criterion), and more recent approaches tailored to specific problems.
Normal Equation and Stacked Form for Tikhonov Regularization

How to compute the Tikhonov regularized solution?

- by using the SVD and spectral filtering:

$$f_{TIKH} = \sum_{i=1}^{\min(m,n)} \frac{\sigma_i (u_i^T y^\delta)}{\sigma_i^2 + \alpha^2} v_i.$$ 

- via the normal equations associated to the minimization problem:

$$f_{TIKH} = (K^T K + \alpha I)^{-1} K^T y^\delta.$$ 

This can be proved by multivariable calculus techniques, and is computationally preferrable.

- stacked form: denote by  $$\tilde{K} = \begin{bmatrix} K \\ \sqrt{\alpha} I \end{bmatrix}$$ and $$\tilde{y}^\delta = \begin{bmatrix} y^\delta \\ 0 \end{bmatrix};$$ then it holds:

$$\tilde{K}^T \tilde{K} f_{TIKH} = \tilde{K}^T \tilde{y}^\delta.$$
Regularization Theory

Tikhonov is just one technique afferent the famous theory of Regularization of inverse problems, which is based on:
Tikhonov regularization

Variational regularization

Extra: deconvolution and denoising problem

Conclusions

Regularization Theory

Tikhonov is just one technique afferent the famous theory of Regularization of inverse problems, which is based on:

1. a family of **regularization functionals** $R_\alpha : \mathbb{R}^m \to \mathbb{R}^n$ varying with $\alpha > 0$:

   $$R_\alpha y \to K^\dagger y \quad \text{as } \alpha \to 0, \quad \forall y \in \mathbb{R}^m.$$

   We only focus on **linear** regularization functionals, which can therefore be represented by a matrix in $\mathbb{R}^{m \times n}$. 

Luca Ratti

Tomography and Regularization

Inverse Problems
Regularization Theory

Tikhonov is just one technique afferent the famous theory of Regularization of inverse problems, which is based on:

1. a family of regularization functionals \( R_\alpha : \mathbb{R}^m \rightarrow \mathbb{R}^n \) varying with \( \alpha > 0 \):

   \[
   R_\alpha y \rightarrow K^\dagger y \quad \text{as} \quad \alpha \rightarrow 0, \quad \forall y \in \mathbb{R}^m.
   \]

   We only focus on linear regularization functionals, which can therefore be represented by a matrix in \( \mathbb{R}^{m \times n} \).

2. a suitable parameter choice rule \( \alpha = \alpha(\delta) \) ensuring

   \[
   R_{\alpha(\delta)} y^\delta \rightarrow K^\dagger y \quad \text{as} \quad \delta \rightarrow 0, \quad \forall y^\delta : \|y^\delta - y\| \leq \delta, \quad \forall y \in \mathbb{R}^m
   \]
Regularization Theory

Tikhonov is just one technique afferent the famous theory of Regularization of inverse problems, which is based on:

1. a family of **regularization functionals** \( R_\alpha : \mathbb{R}^m \to \mathbb{R}^n \) varying with \( \alpha > 0 \):

\[
R_\alpha y \to K^\dagger y \quad \text{as} \quad \alpha \to 0, \quad \forall y \in \mathbb{R}^m.
\]

We only focus on **linear** regularization functionals, which can therefore be represented by a matrix in \( \mathbb{R}^{m \times n} \).

2. a suitable **parameter choice rule** \( \alpha = \alpha(\delta) \) ensuring

\[
R_{\alpha(\delta)} y^{\delta} \to K^\dagger y \quad \text{as} \quad \delta \to 0, \quad \forall y^{\delta} : \| y^{\delta} - y \| \leq \delta, \quad \forall y \in \mathbb{R}^m
\]

A desirable property is that \( R_\alpha \) are **more stable** than \( K^\dagger \):

\[
\| R_\alpha \| \leq \| K^\dagger \| \quad \forall \alpha > 0
\]
Tikhonov regularization

Variational regularization

Extra: deconvolution and denoising problem

Conclusions

Regularization Theory

Tikhonov is just one technique afferent the famous theory of Regularization of inverse problems, which is based on:

1. a family of **regularization functionals** $R_\alpha : \mathbb{R}^m \to \mathbb{R}^n$ varying with $\alpha > 0$:

   $$R_\alpha y \to K^\dagger y \quad \text{as} \quad \alpha \to 0, \quad \forall y \in \mathbb{R}^m.$$

   We only focus on **linear** regularization functionals, which can therefore be represented by a matrix in $\mathbb{R}^{m \times n}$.

2. a suitable **parameter choice rule** $\alpha = \alpha(\delta)$ ensuring

   $$R_{\alpha(\delta)} y^\delta \to K^\dagger y \quad \text{as} \quad \delta \to 0, \quad \forall y^\delta : \|y^\delta - y\| \leq \delta, \quad \forall y \in \mathbb{R}^m.$$

   A desirable property is that $R_\alpha$ are **more stable** than $K^\dagger$:

   $$\|R_\alpha\| \leq \|K^\dagger\| \quad \forall \alpha > 0$$

Remark: a parameter choice $\alpha = \alpha(\delta)$ is defined **a priori**, since it holds for any $y$ and any perturbation $y^\delta$. We will focus more on **a posteriori** (heuristic) rules $\alpha = \alpha(\delta, y^\delta)$. 
Graphical interpretation

What happens if we compute $f_{\alpha,\delta} = R_{\alpha} y^{\delta}$ instead of $f_{\delta} = K^{\dagger} y^{\delta}$?

Error bound for the approximation of $f = f^{\dagger} = K^{\dagger} y$

By using $f_{\delta}$: $\| f^{\dagger} - f_{\delta} \| = \| K^{\dagger} y - K^{\dagger} y^{\delta} \| \leq \| K^{\dagger} \| \| \delta \|$

By using $f_{\alpha,\delta}$: $\| f^{\dagger} - f_{\alpha,\delta} \| \leq \| f^{\dagger} - R_{\alpha} y \| + \| R_{\alpha} \| \| \delta \|$. 
Graphical interpretation

What happens if we compute \( f_{\alpha,\delta} = R\alpha y^\delta \) instead of \( f_\delta = K^\dagger y^\delta \)?

**Error bound for the approximation of** \( f = f^\dagger = K^\dagger y \)

By using \( f_\delta \):
\[
\| f^\dagger - f_\delta \| = \| K^\dagger y - K^\dagger y^\delta \| \leq \| K^\dagger \| \delta
\]

By using \( f_{\alpha,\delta} \):
\[
\| f^\dagger - f_{\alpha,\delta} \| \leq \| f^\dagger - R\alpha y \| + \| R\alpha \| \delta.
\]
**Graphical interpretation**

What happens if we compute \( f_{\alpha,\delta} = R_\alpha y^\delta \) instead of \( f_\delta = K^\dagger y^\delta \)?

**Error bound for the approximation of \( f = f^\dagger = K^\dagger y \)**

By using \( f_\delta \):

\[
\| f^\dagger - f_\delta \| = \| K^\dagger y - K^\dagger y^\delta \| \leq \| K^\dagger \| \delta
\]

By using \( f_{\alpha,\delta} \):

\[
\| f^\dagger - f_{\alpha,\delta} \| \leq \| f^\dagger - R_\alpha y \| + \| R_\alpha \| \delta.
\]
Graphical interpretation

What happens if we compute \( f_{\alpha,\delta} = R_\alpha y^\delta \) instead of \( f_\delta = K^{\dagger} y^\delta \)?

Error bound for the approximation of \( f = f^{\dagger} = K^{\dagger} y \)

By using \( f_\delta \): \( \| f^{\dagger} - f_\delta \| = \| K^{\dagger} y - K^{\dagger} y^\delta \| \leq \| K^{\dagger} \| \delta \)

By using \( f_{\alpha,\delta} \): \( \| f^{\dagger} - f_{\alpha,\delta} \| \leq \| f^{\dagger} - R_\alpha y \| + \| R_\alpha \| \delta \).
Graphical interpretation

What happens if we compute $f_{\alpha,\delta} = R_{\alpha}y^\delta$ instead of $f_{\delta} = K^\dagger y^\delta$?

Error bound for the approximation of $f = f^\dagger = K^\dagger y$

By using $f_{\delta}$: $\|f^\dagger - f_{\delta}\| = \|K^\dagger y - K^\dagger y^\delta\| \leq \|K^\dagger\| \delta$

By using $f_{\alpha,\delta}$: $\|f^\dagger - f_{\alpha,\delta}\| \leq \|f^\dagger - R_{\alpha}y\| + \|R_{\alpha}\| \delta$. 

![Graphical interpretation](image-url)
Graphical interpretation

What happens if we compute $f_{\alpha,\delta} = R_{\alpha} y^\delta$ instead of $f_{\delta} = K^\dagger y^\delta$?

Error bound for the approximation of $f = f^\dagger = K^\dagger y$

By using $f_{\delta}$: $\|f^\dagger - f_{\delta}\| = \|K^\dagger y - K^\dagger y^\delta\| \leq \|K^\dagger\| \delta$

By using $f_{\alpha,\delta}$: $\|f^\dagger - f_{\alpha,\delta}\| \leq \|f^\dagger - R_{\alpha} y\| + \|R_{\alpha}\| \delta$. 
Graphical interpretation

What happens if we compute \( f_{\alpha,\delta} = R_{\alpha} y^{\delta} \) instead of \( f_{\delta} = K^\dagger y^{\delta} \)?

Error bound for the approximation of \( f = f^\dagger = K^\dagger y \)

By using \( f_{\delta} \):
\[
\| f^\dagger - f_{\delta} \| = \| K^\dagger y - K^\dagger y^{\delta} \| \leq \| K^\dagger \| \delta
\]

By using \( f_{\alpha,\delta} \):
\[
\| f^\dagger - f_{\alpha,\delta} \| \leq \| f^\dagger - R_{\alpha} y \| + \| R_{\alpha} \| \delta.
\]
What happens if we compute $f_{\alpha, \delta} = R_{\alpha} y^\delta$ instead of $f_{\delta} = K^\dagger y^\delta$?

### Error bound for the approximation of $f = f^\dagger = K^\dagger y$

By using $f_{\delta}$: $\| f^\dagger - f_{\delta} \| = \| K^\dagger y - K^\dagger y^\delta \| \leq \| K^\dagger \| \delta$

By using $f_{\alpha, \delta}$: $\| f^\dagger - f_{\alpha, \delta} \| \leq \| f^\dagger - R_{\alpha} y \| + \| R_{\alpha} \| \delta$. 

![Graphical interpretation](image.png)
Graphical interpretation gets complicated

The previous pictures showed a single $\alpha$ at a time. Let’s visualize the error bound as a function of $\delta$ and $\alpha$: 
Graphical interpretation gets complicated

The previous pictures showed a single $\alpha$ at a time. Let’s visualize the error bound as a function of $\delta$ and $\alpha$:
Graphical interpretation gets complicated

The previous pictures showed a single $\alpha$ at a time. Let's visualize the error bound as a function of $\delta$ and $\alpha$: 
Graphical interpretation gets complicated

The previous pictures showed a single $\alpha$ at a time. Let's visualize the error bound as a function of $\delta$ and $\alpha$: 

![Graphical representation showing the error bound as a function of $\delta$ and $\alpha$.]
Graphical interpretation gets complicated

The previous pictures showed a single $\alpha$ at a time. Let's visualize the error bound as a function of $\delta$ and $\alpha$: 

![Graphical representation of error bound as a function of $\delta$ and $\alpha$.]
Graphical interpretation gets complicated

The previous pictures showed a single $\alpha$ at a time. Let's visualize the error bound as a function of $\delta$ and $\alpha$: 

![Graphical representation]

$Luca\ Ratti$  
$Tomography\ and\ Regularization$  
$Inverse\ Problems$
The previous pictures showed a single $\alpha$ at a time. Let's visualize the error bound as a function of $\delta$ and $\alpha$:
Graphical interpretation gets complicated

The previous pictures showed a single $\alpha$ at a time. Let's visualize the error bound as a function of $\delta$ and $\alpha$: 
Graphical interpretation gets complicated

The previous pictures showed a single $\alpha$ at a time. Let's visualize the error bound as a function of $\delta$ and $\alpha$:
The parameter choice rule is a relationship $\alpha = \alpha(\delta)$ ensuring that, along the curve $\|f^\dagger - R_{\alpha(\delta)}y^\delta\|$, we are converging to 0.
The parameter choice rule is a relationship $\alpha = \alpha(\delta)$ ensuring that, along the curve $\| f^\dagger - R_{\alpha(\delta)} y^\delta \|$, we are converging to 0.
Graphical interpretation - parameter choice rule

The parameter choice rule is a relationship $\alpha = \alpha(\delta)$ ensuring that, along the curve $\| f^\dagger - R_{\alpha(\delta)}y^\delta \|$, we are converging to 0.
Variational Regularization

One of the most broad family of regularization methods is the Variational Regularization. General expression:

\[
R_\alpha(y^\delta) = \arg\min_f \left\{ \frac{1}{2} \left\| Kf - y^\delta \right\|^2 + \alpha \mathcal{R}(f) \right\}
\]

- The data fidelity (or data fitting) term \( \left\| Kf - y^\delta \right\|^2 \) keeps the estimation of the solution close to the data under the forward physical system.
- \( \mathcal{R}(f) \) encodes some a priori knowledge about the solution that we want to promote in the regularized solution.
- The regularization parameter \( \alpha > 0 \) controls the trade-off between a good fit and the requirements from the regularization.

Most important examples:

- Tikhonov regularization: \( \mathcal{R}(f) = \| f \|^2 \). A priori knowledge: the solution has small norm.
- Generalized Tikhonov regularization: \( \mathcal{R}(f) = \| L(f - f^*) \|^2 \).
Generalized Tikhonov Regularization

Via GTIKH, it is possible to encode different \textit{a priori} information about the solution of the inverse problem. For instance:

- $f$ is close to a known $f^*$

$$f^\text{GTIKH} = \arg \min_f \left\{ \|Kf - y\|^2 + \alpha \|f - f^*\|^2 \right\}$$

- $f$ is known to be smooth

$$f^\text{GTIKH} = \arg \min_f \left\{ \|Kf - y\|^2 + \alpha \|Lf\|^2 \right\},$$

where $L$ is the discretization of a differential operator.

- $f$ has similar smoothing properties as $f^*$

$$f^\text{GTIKH} = \arg \min_f \left\{ \|Kf - y\|^2 + \alpha \|L(f - f^*)\|^2 \right\}$$
Sparsity-promoting regularization

What happens if we select \( R(f) = \| f \|_1 = \sum_{i=1}^{n} |f_i| \) ?

\[
\arg\min_f \left\{ \frac{1}{2} \| Kf - y^\delta \|_2^2 + \alpha \| f \|_1 \right\}.
\]

Under some theoretical assumptions, we can show that this allows to enforce that the regularized solution is \textit{sparse}, i.e., only few components are different from 0. Intuition:
Sparsity-promoting regularization

What happens if we select $\mathcal{R}(f) = \|f\|_1 = \sum_{i=1}^{n} |f_i|$?

$$\arg\min_f \left\{ \frac{1}{2} \|Kf - y^\delta\|_2^2 + \alpha \|f\|_1 \right\}.$$ 

Under some theoretical assumptions, we can show that this allows to enforce that the regularized solution is **sparse**, i.e., only few components are different from 0. Intuition:

$$|x_1|^2 + |x_2|^2 = \text{const} \quad \text{and} \quad |x_1| + |x_2| = \text{const}.$$
Sparsity-promoting regularization: extensions

- sparsity with respect to a different basis. Let $W$ be a matrix representing the change into a different basis on $\mathbb{R}^n$. Then, the variational formulation

$$f^W = \arg\min_f \left\{ \frac{1}{2} \| Kf - y^\delta \|_2^2 + \alpha \| Wf \|_1 \right\}$$

promotes sparsity wrt the new basis. Typical application: $W$ is a wavelet basis.

- Total Variation. Let $L$ be the discretization of the first derivative operator:

$$f^{TV} = \arg\min_f \left\{ \frac{1}{2} \| Kf - y^\delta \|_2^2 + \alpha \| Lf \|_1 \right\}$$

is called Total Variation.

Total Variation (TV) regularization promotes sparsity in the derivative, in other words favouring piece-wise constantness.
Naive Reconstruction (Moore-Penrose Pseudoinverse)

Original phantom

$f^\dagger$: RE = 100%
Truncated SVD Regularization

Original phantom

\( f^{TSVD} \): RE = 35%
Tikhonov Regularization

Original phantom

\( f^{\text{TIKH}} : \text{RE} = 32\% \)
Nonnegativity Constrained Wavelet-based Regularization

Original phantom

$f^{WLET}_+$: RE = 26%
Nonnegativity Constrained Total Variation Regularization

Original phantom

\( f^+_{\text{TV}} : \text{RE} = 3\% \)
Pictures can be represented by vectors

Every pixel of a (greyscale) picture can be represented by a number between 0 (black) and 1 (white):
Pictures can be represented by vectors

Every pixel of a (greyscale) picture can be represented by a number between 0 (black) and 1 (white):

\[
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\
0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\
0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]
Pictures can be represented by vectors

Every pixel of a (greyscale) picture can be represented by a number between 0 (black) and 1 (white):

\[
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.9 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0.8 & 1 & 1 & 1 & 1 & 1 & 0.8 & 0 & 0 \\
0 & 0.4 & 1 & 0.2 & 0.1 & 1 & 0.2 & 0.1 & 1 & 0.4 & 0 \\
0.1 & 1 & 1 & 0.1 & 0 & 1 & 0.1 & 0 & 1 & 1 & 0 \\
0.2 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\
0.1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\
0 & 1 & 0.9 & 0 & 1 & 1 & 1 & 0 & 0.9 & 1 & 0 \\
0 & 0 & 1 & 0.9 & 0 & 0 & 0 & 0 & 0.9 & 1 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]
Pictures can be represented by vectors

Every pixel of a (greyscale) picture can be represented by a number between 0 (black) and 1 (white):

256x256
Convolution is a linear operation on pictures

**Convolution** is a phenomenon occurring in many physical contexts. In particular, convolution (or blurring) occurs when we take pictures without adjusting the focus of the camera.
Convolution is a linear operation on pictures

**Convolution** is a phenomenon occurring in many physical contexts. In particular, convolution (or blurring) occurs when we take pictures without adjusting the focus of the camera.

Physical intuition: light rays are not properly collected by the lens!
Convolution is a linear operation on pictures

**Convolution** is a phenomenon occurring in many physical contexts. In particular, convolution (or blurring) occurs when we take pictures without adjusting the focus of the camera.

Mathematical intuition: each pixel of the blurred picture is a (weighted) average of the neighboring pixels in the original picture.

We can interpret this as a linear operation on the elements of the original picture.
Deblurring (and denoising): a linear inverse problem

- Consider a vector $f \in \mathbb{R}^n$ representing a picture ($n$: the number of pixels).
- Represent the effect of blurring (linear operation) as a matrix $K : \mathbb{R}^n \rightarrow \mathbb{R}^n$. The blurred version of $f$ is thus $y = Kf$.
- Suppose some additional noise is present on the picture $y^\delta : \|y - y^\delta\| \leq \delta$. 

Luca Ratti  
Tomography and Regularization
Deblurring (and denoising): a linear inverse problem

- Consider a vector \( f \in \mathbb{R}^n \) representing a picture (\( n \): the number of pixels).
- Represent the effect of blurring (linear operation) as a matrix \( K : \mathbb{R}^n \rightarrow \mathbb{R}^n \). The blurred version of \( f \) is thus \( y = Kf \).
- Suppose some additional noise is present on the picture \( y^\delta \): \( \| y - y^\delta \| \leq \delta \).
Deblurring (and denoising): a linear inverse problem

- Consider a vector $f \in \mathbb{R}^n$ representing a picture ($n$: the number of pixels).
- Represent the effect of blurring (linear operation) as a matrix $K : \mathbb{R}^n \to \mathbb{R}^n$. The blurred version of $f$ is thus $y = Kf$.
- Suppose some additional noise is present on the picture $y^\delta : \|y - y^\delta\| \leq \delta$. 

[Image of blurred and noisy data]
Deblurring (and denoising): a linear inverse problem

- Consider a vector \( f \in \mathbb{R}^n \) representing a picture (\( n \): the number of pixels).
- Represent the effect of blurring (linear operation) as a matrix \( K : \mathbb{R}^n \rightarrow \mathbb{R}^n \). The blurred version of \( f \) is thus \( y = Kf \).
- Suppose some additional noise is present on the picture \( y^\delta : \|y - y^\delta\| \leq \delta \).

Inverse problem (deconvolution and denoising)

Given \( y^\delta \in \mathbb{R}^n \) s.t. \( \|y^\delta - Kf\| \leq \delta \), find \( f \in \mathbb{R}^n \).
The importance of regularization

No regularization

Tikhonov, $\alpha = 0.15$

Naive inversion vs Tikhonov regularization.
The importance of regularization

Generalized Tikhonov, $\alpha = 0.15$

Generalized Tikhonov, $\alpha = 1.75$

Generalized Tikhonov regularization: the effect of different values of $\alpha$. 
The importance of regularization

Total Variation regularization: the effect of different values of $\alpha$. 
What regularization strategy would you pick? - I

Exercise 1

1. We are given a blurred and noise picture to deconvolve (filter: 5x5, Gaussian).

2. According to a priori knowledge, what regularization strategy would you recommend?
What regularization strategy would you pick? - 1

Exercise 1

1. We are given a blurred and noise picture to deconvolve (filter: 5x5, Gaussian).

2. According to a priori knowledge, what regularization strategy would you recommend?
Exercise 1

1. We are given a blurred and noise picture to deconvolve (filter: 5x5, Gaussian).
2. According to a priori knowledge, what regularization strategy would you recommend?

Hint: Leonardo Da Vinci was an expert of **sfumato**, a painting technique consisting in softening the transition between colours.
What regularization strategy would you pick? - I

Original Image

Blurred and noisy data
What regularization strategy would you pick? - I

Original Image

Tikhonov, $\alpha = 0.15$
What regularization strategy would you pick? - I

Original Image

Generalized Tikhonov, $\alpha = 0.9$
What regularization strategy would you pick? - 1

Original Image

TV, $\alpha = 0.125$
Exercise 2

1. We are given a blurred and noise picture to deconvolve (filter: 5x5, Gaussian).

2. According to a priori knowledge, what regularization strategy would you recommend?
What regularization strategy would you pick? - I

Exercise 2

1. We are given a blurred and noise picture to deconvolve (filter: 5x5, Gaussian).
2. According to a priori knowledge, what regularization strategy would you recommend?

Hint: Piet Mondrian was the founder of Neoplasticism, whose main elements are the use of primary colors (red, blue and yellow), three primary values (black, white and gray) and two primary directions (horizontal and vertical).
What regularization strategy would you pick? - I

Exercise 2

1. We are given a blurred and noise picture to deconvolve (filter: 5x5, Gaussian).
2. According to a priori knowledge, what regularization strategy would you recommend?

Hint: Piet Mondrian was the founder of Neoplasticism, whose main elements are the use of primary colors (red, blue and yellow), three primary values (black, white and gray) and two primary directions (horizontal and vertical).
What regularization strategy would you pick? - I
What regularization strategy would you pick? - I

Original Image

Tikhonov, $\alpha = 0.15$
What regularization strategy would you pick? - I

Original Image

Generalized Tikhonov, $\alpha = 0.9$
What regularization strategy would you pick? - I

Original Image

TV, $\alpha = 0.125$
Conclusions

What we saw:

- regularization is a power tool for recovering well-posedness in inverse problems
- the choice of the regularization strategy (Tikhonov, generalized Tikhonov, sparsity-promoting, TV, ...) is crucial and depends on prior information we have on the solution
- the choice of the regularization parameter must be done carefully

What we did not see (and could be a nice extension):

- heuristic rules for the choice of $\alpha$
- how to implement $\ell^1$ minimization and $TV$
- how to introduce constraint in the variational regularization problem
Some References

Inverse Problems:

- Vogel, *Computational methods for inverse problems*, 2002
- Hansen, *Discrete inverse problems*, 2010

X-ray Tomography:

- Natterer, *The mathematics of computerized tomography*, 1986
- Buzug: *Computed Tomography: From Photon Statistics to Modern Cone-Beam CT*, 2008
- Epstein, *Introduction to the mathematics of medical imaging*, 2008
Some References

- **Total variation and wavelets**
  - Mallat, *A wavelet tour of signal processing*, 1999

- **Optimization**
  - Boyd & Vandenberghe, *Convex optimization*, 2004
  - Rockafellar, *Convex optimization*, 1996
  - Daubechies, Defrise & De Mol, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*, 2004