Math 4/896: Seminar in Mathematics
Topic: Inverse Theory

Instructor: Thomas Shores
Department of Mathematics

Lecture 1, January 10, 2006
AvH 10

Outline

Basics

Regularization:

This means "turn an ill-posed problem into a well-posed 'near by' problem". Most common method is Tikhonov regularization, which is motivated in context of our possibly ill-posed $G\mathbf{m} = \mathbf{d}$, i.e., minimize $\|G\mathbf{m} - \mathbf{d}\|_2$, problem by:

- Problem: minimize $\|\mathbf{m}\|_2$ subject to $\|\mathbf{Gm} - \mathbf{d}\|_2 \leq \delta$
- Problem: minimize $\|G\mathbf{m} - \mathbf{d}\|_2$ subject to $\|\mathbf{m}\|_2 \leq \epsilon$
- Problem: **(damped least squares)** minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha^2 \|\mathbf{m}\|_2^2$. This is the **Tikhonov regularization** of the original problem.
- Problem: find minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \leq c$.e function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \geq 0$.

Basics

Regularization:

This means "turn an ill-posed problem into a well-posed 'near by' problem". Most common method is Tikhonov regularization, which is motivated in context of our possibly ill-posed $G\mathbf{m} = \mathbf{d}$, i.e., minimize $\|G\mathbf{m} - \mathbf{d}\|_2$, problem by:

- Problem: minimize $\|\mathbf{m}\|_2$ subject to $\|\mathbf{Gm} - \mathbf{d}\|_2 \leq \delta$
- Problem: minimize $\|G\mathbf{m} - \mathbf{d}\|_2$ subject to $\|\mathbf{m}\|_2 \leq \epsilon$
- Problem: (**damped least squares**) minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha^2 \|\mathbf{m}\|_2^2$. This is the **Tikhonov regularization** of the original problem.
- Problem: find minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \leq c$.e function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \geq 0$.

Basics

Regularization:

This means “turn an ill-posed problem into a well-posed 'near by' problem”. Most common method is Tikhonov regularization, which is motivated in context of our possibly ill-posed $G\mathbf{m} = \mathbf{d}$, i.e., minimize $\|G\mathbf{m} - \mathbf{d}\|_2$, problem by:

- Problem: minimize $\|\mathbf{m}\|_2$ subject to $\|\mathbf{Gm} - \mathbf{d}\|_2 \leq \delta$
- Problem: minimize $\|G\mathbf{m} - \mathbf{d}\|_2$ subject to $\|\mathbf{m}\|_2 \leq \epsilon$
- Problem: (**damped least squares**) minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha^2 \|\mathbf{m}\|_2^2$. This is the **Tikhonov regularization** of the original problem.
- Problem: find minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \leq c$.e function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \geq 0$.

Basics

Regularization:

This means "turn an ill-posed problem into a well-posed 'near by' problem". Most common method is Tikhonov regularization, which is motivated in context of our possibly ill-posed $G\mathbf{m} = \mathbf{d}$, i.e., minimize $\|G\mathbf{m} - \mathbf{d}\|_2$, problem by:

- Problem: minimize $\|\mathbf{m}\|_2$ subject to $\|\mathbf{Gm} - \mathbf{d}\|_\mathbf{2} \leq \delta$
- Problem: minimize $\|G\mathbf{m} - \mathbf{d}\|_2$ subject to $\|\mathbf{m}\|_2 \leq \epsilon$
- Problem: **(damped least squares)** minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha^2 \|\mathbf{m}\|_2^2$. This is the **Tikhonov regularization** of the original problem.
- Problem: find minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \leq c$.e function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \geq 0$.

Basics

Regularization:

This means "turn an ill-posed problem into a well-posed 'near by' problem". Most common method is Tikhonov regularization, which is motivated in context of our possibly ill-posed $G\mathbf{m} = \mathbf{d}$, i.e., minimize $\|G\mathbf{m} - \mathbf{d}\|_2$, problem by:

- Problem: minimize $\|\mathbf{m}\|_2$ subject to $\|\mathbf{Gm} - \mathbf{d}\|_2 \leq \delta$
- Problem: minimize $\|G\mathbf{m} - \mathbf{d}\|_2$ subject to $\|\mathbf{m}\|_2 \leq \epsilon$
- Problem: (**damped least squares**) minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha^2 \|\mathbf{m}\|_2^2$. This is the **Tikhonov regularization** of the original problem.
- Problem: find minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \leq c$.e function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \geq 0$.

Basics

Regularization:

All of the above problems are equivalent under mild restrictions thanks to the principle of Lagrange multipliers:

- Minima of $f(\mathbf{x})$ occur at **stationary points** of $f(x)$ ($\nabla f = 0$.)
- Minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \leq c$ must occur at stationary points of function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \geq 0$ (we can write $\lambda = \alpha^2$ to emphasize non-negativity.)
- We can see why this is true in the case of a two dimensional $\mathbf{x}$ by examining contour curves.
- Square the terms in the first two problems and we see that the associated Lagrangians are related if we take reciprocals of α.
- Various values of α give a trade-off between the instability of the unmodified least squares problem and loss of accuracy of the smoothed problem. This can be understood by tracking the value of the minimized function in the form of a path depending on δ, ϵ or α.

Basics

Regularization:

All of the above problems are equivalent under mild restrictions thanks to the principle of Lagrange multipliers:

- Minima of $f(\mathbf{x})$ occur at **stationary points** of $f(x)$ ($\nabla f = 0$.)
- Minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \leq c$ must occur at stationary points of function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \geq 0$ (we can write $\lambda = \alpha^2$ to emphasize non-negativity.)
- We can see why this is true in the case of a two dimensional $\mathbf{x}$ by examining contour curves.
- Square the terms in the first two problems and we see that the associated Lagrangians are related if we take reciprocals of α.
- Various values of α give a trade-off between the instability of the unmodified least squares problem and loss of accuracy of the smoothed problem. This can be understood by tracking the value of the minimized function in the form of a path depending on δ, ϵ or α.

Basics

Regularization:

All of the above problems are equivalent under mild restrictions thanks to the principle of Lagrange multipliers:

- Minima of $f(\mathbf{x})$ occur at **stationary points** of $f(x)$ ($\nabla f = 0$.)
- Minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \leq c$ must occur at stationary points of function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \geq 0$ (we can write $\lambda = \alpha^2$ to emphasize non-negativity.)
- We can see why this is true in the case of a two dimensional $\mathbf{x}$ by examining contour curves.
- Square the terms in the first two problems and we see that the associated Lagrangians are related if we take reciprocals of α.
- Various values of α give a trade-off between the instability of the unmodified least squares problem and loss of accuracy of the smoothed problem. This can be understood by tracking the value of the minimized function in the form of a path depending on δ, ϵ or α.

Basics

Regularization:

All of the above problems are equivalent under mild restrictions thanks to the principle of Lagrange multipliers:

- Minima of $f(\mathbf{x})$ occur at **stationary points** of $f(x)$ ($\nabla f = 0$.)
- Minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \leq c$ must occur at stationary points of function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \geq 0$ (we can write $\lambda = \alpha^2$ to emphasize non-negativity.)
- We can see why this is true in the case of a two dimensional $\mathbf{x}$ by examining contour curves.
- Square the terms in the first two problems and we see that the associated Lagrangians are related if we take reciprocals of α.
- Various values of α give a trade-off between the instability of the unmodified least squares problem and loss of accuracy of the smoothed problem. This can be understood by tracking the value of the minimized function in the form of a path depending on δ, ϵ or α.

Basics

Regularization:

All of the above problems are equivalent under mild restrictions thanks to the principle of Lagrange multipliers:

- Minima of $f(\mathbf{x})$ occur at **stationary points** of $f(x)$ $(\nabla f = 0.)$
- Minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \le c$ must occur at stationary points of function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \ge 0$ (we can write $\lambda = \alpha^2$ to emphasize non-negativity.)
- We can see why this is true in the case of a two dimensional $\mathbf{x}$ by examining contour curves.
- Square the terms in the first two problems and we see that the associated Lagrangians are related if we take reciprocals of α.
- Various values of α give a trade-off between the instability of the unmodified least squares problem and loss of accuracy of the smoothed problem. This can be understood by tracking the value of the minimized function in the form of a path depending on δ, ϵ or α.

Regularization:

All of the above problems are equivalent under mild restrictions thanks to the principle of Lagrange multipliers:

- Minima of $f(\mathbf{x})$ occur at **stationary points** of $f(x)$ $(\nabla f = 0.)$
- Minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \leq c$ must occur at stationary points of function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \geq 0$ (we can write $\lambda = \alpha^2$ to emphasize non-negativity.)
- We can see why this is true in the case of a two dimensional $\mathbf{x}$ by examining contour curves.
- Square the terms in the first two problems and we see that the associated Lagrangians are related if we take reciprocals of α.
- Various values of α give a trade-off between the instability of the unmodified least squares problem and loss of accuracy of the smoothed problem. This can be understood by tracking the value of the minimized function in the form of a path depending on δ, ϵ or α.

Midterm Review

Rules of the Game:

- The midterm will have two parts.
- An in-class part that is closed book, closed notes, NO calculators, laptops, cell phones, blackberries, etc., etc. This part is worth 80 points. This exam will be administered on Thursday, March 9.
- A take-home part that is worth 50 points. This exam will be available on the class home page Tuesday morning. It will be due on Friday, March 10, 5:00 pm. It can either be hardcopy or in the form of a pdf (or even Microsoft Word) file which you can email to me. You must show all your work, including copies of any scripts that you used and their relevant outputs.

Midterm Review

Rules of the Game:

- The midterm will have two parts.
- An in-class part that is closed book, closed notes, NO calculators, laptops, cell phones, blackberries, etc., etc. This part is worth 80 points. This exam will be administered on Thursday, March 9.
- A take-home part that is worth 50 points. This exam will be available on the class home page Tuesday morning. It will be due on Friday, March 10, 5:00 pm. It can either be hardcopy or in the form of a pdf (or even Microsoft Word) file which you can email to me. You must show all your work, including copies of any scripts that you used and their relevant outputs.

Midterm Review

Rules of the Game:

- The midterm will have two parts.
- An in-class part that is closed book, closed notes, NO calculators, laptops, cell phones, blackberries, etc., etc. This part is worth 80 points. This exam will be administered on Thursday, March 9.
- A take-home part that is worth 50 points. This exam will be available on the class home page Tuesday morning. It will be due on Friday, March 10, 5:00 pm. It can either be hardcopy or in the form of a pdf (or even Microsoft Word) file which you can email to me. You must show all your work, including copies of any scripts that you used and their relevant outputs.

Midterm Review

Rules of the Game:

- The midterm will have two parts.
- An in-class part that is closed book, closed notes, NO calculators, laptops, cell phones, blackberries, etc., etc. This part is worth 80 points. This exam will be administered on Thursday, March 9.
- A take-home part that is worth 50 points. This exam will be available on the class home page Tuesday morning. It will be due on Friday, March 10, 5:00 pm. It can either be hardcopy or in the form of a pdf (or even Microsoft Word) file which you can email to me. You must show all your work, including copies of any scripts that you used and their relevant outputs.

Midterm Review

More Rules of the Game:

- There is to be absolutely no consultation of any kind with anyone else other than me about the exam. If there are points of clarification or (heavens forbid!) errors, I will post them on our message board.
- ALL materials used in your work that have not been provided by me for this course must be explicitly credited in your write-up.

Midterm Review

More Rules of the Game:

- There is to be absolutely no consultation of any kind with anyone else other than me about the exam. If there are points of clarification or (heavens forbid!) errors, I will post them on our message board.
- ALL materials used in your work that have not been provided by me for this course must be explicitly credited in your write-up.

Midterm Review

More Rules of the Game:

- There is to be absolutely no consultation of any kind with anyone else other than me about the exam. If there are points of clarification or (heavens forbid!) errors, I will post them on our message board.
- ALL materials used in your work that have not been provided by me for this course must be explicitly credited in your write-up.

Midterm Review

Material Covered in Midterm:

- All lecture notes material covering ABT text Chapters 1-4 inclusive.
- All lecture notes material covering ABT text Chapter 5, Section 1.
- Excludes all lecture notes material covering our intro/review of Matlab, linear algebra and probability/statistics *per se*.
- "Per se" does NOT exclude linear algebra/probability/stat topics that are a part of the lecture notes on ABT text.

Midterm Review

Material Covered in Midterm:

- All lecture notes material covering ABT text Chapters 1-4 inclusive.
- All lecture notes material covering ABT text Chapter 5, Section 1.
- Excludes all lecture notes material covering our intro/review of Matlab, linear algebra and probability/statistics *per se*.
- "Per se" does NOT exclude linear algebra/probability/stat topics that are a part of the lecture notes on ABT text.

Midterm Review

Material Covered in Midterm:

- All lecture notes material covering ABT text Chapters 1-4 inclusive.
- All lecture notes material covering ABT text Chapter 5, Section 1.
- Excludes all lecture notes material covering our intro/review of Matlab, linear algebra and probability/statistics *per se*.
- "Per se" does NOT exclude linear algebra/probability/stat topics that are a part of the lecture notes on ABT text.

Midterm Review

Material Covered in Midterm:

- All lecture notes material covering ABT text Chapters 1-4 inclusive.
- All lecture notes material covering ABT text Chapter 5, Section 1.
- Excludes all lecture notes material covering our intro/review of Matlab, linear algebra and probability/statistics *per se*.
- “Per se” does NOT exclude linear algebra/probability/stat topics that are a part of the lecture notes on ABT text.

Midterm Review

Material Covered in Midterm:

- All lecture notes material covering ABT text Chapters 1-4 inclusive.
- All lecture notes material covering ABT text Chapter 5, Section 1.
- Excludes all lecture notes material covering our intro/review of Matlab, linear algebra and probability/statistics *per se*.
- "Per se" does NOT exclude linear algebra/probability/stat topics that are a part of the lecture notes on ABT text.

Sample Questions

Sample Questions:

- Inverse problems are typically ill-posed. What does that mean? Give a specific example of an ill-posed problem.
- Find the normal equations for the system

$$\begin{aligned} -2x_1 + 2x_2 &= 1 \\ -x_1 + x_2 &= 1 \\ x_1 + x_2 &= 2. \end{aligned}$$

  What is the residual for this problem?
- What is the difference between a rank deficient problem and and ill-posed problem $G\mathbf{m} = \mathbf{d}$?

Sample Questions

Sample Questions:

- Inverse problems are typically ill-posed. What does that mean? Give a specific example of an ill-posed problem.
- Find the normal equations for the system

$$\begin{aligned} -2x_1 + 2x_2 &= 1 \\ -x_1 + x_2 &= 1 \\ x_1 + x_2 &= 2. \end{aligned}$$

  What is the residual for this problem?
- What is the difference between a rank deficient problem and and ill-posed problem $G\mathbf{m} = \mathbf{d}$?

Sample Questions

Sample Questions:

- Inverse problems are typically ill-posed. What does that mean? Give a specific example of an ill-posed problem.
- Find the normal equations for the system

$$\begin{aligned} -2x_1 + 2x_2 &= 1 \\ -x_1 + x_2 &= 1 \\ x_1 + x_2 &= 2. \end{aligned}$$

  What is the residual for this problem?
- What is the difference between a rank deficient problem and and ill-posed problem $G\mathbf{m} = \mathbf{d}$?

Sample Questions

Sample Questions:

- Inverse problems are typically ill-posed. What does that mean? Give a specific example of an ill-posed problem.
- Find the normal equations for the system

$$\begin{aligned} -2x_1 + 2x_2 &= 1 \\ -x_1 + x_2 &= 1 \\ x_1 + x_2 &= 2. \end{aligned}$$

  What is the residual for this problem?
- What is the difference between a rank deficient problem and and ill-posed problem $G\mathbf{m} = \mathbf{d}$?

Sample Questions:

- Assume that for a full column rank matrix G, $G^T G$ is invertible and $G\mathbf{m}_{true} = \mathbf{d}_{true}$ is satisfied by a model and exact data. Show that if the coordinates of the error vector $\mathbf{d} - \mathbf{d}_{true}$ are independent and identically distributed with mean zero and standard deviation σ, then $E[\mathbf{m}] = \mathbf{m}_{true}$, where $\mathbf{m}$ solves the least squares problem for $G\mathbf{m} = \mathbf{d}$.
- State the SVD Theorem and explain how it enables us to compute the null space of G.
- Suppose we want to solve the IFK $d(s) = \int_0^1 g(x,s)\, m(x)\, dx$, given two values $d(s_i)$, $i = 1, 2$, for the parameter $m(x)$. Write out the linear system that results if we use collocation at the points $1/4$ and $3/4$ and the midpoint rule for integration.
- What is the Tikhonov regularization of the (possibly) ill-posed problem $G\mathbf{m} = \mathbf{d}$? Exhibit the Lagrangian of the problem of minimizing $\|G\mathbf{m} - \mathbf{d}\|_2$ subject to the constraint $\|\mathbf{m}\|_2 \leq \epsilon$ and show how it is related to Tikhonov regularization.

Sample Questions:

- Assume that for a full column rank matrix G, $G^T G$ is invertible and $G\mathbf{m}_{true} = \mathbf{d}_{true}$ is satisfied by a model and exact data. Show that if the coordinates of the error vector $\mathbf{d} - \mathbf{d}_{true}$ are independent and identically distributed with mean zero and standard deviation σ, then $E[\mathbf{m}] = \mathbf{m}_{true}$, where $\mathbf{m}$ solves the least squares problem for $G\mathbf{m} = \mathbf{d}$.
- State the SVD Theorem and explain how it enables us to compute the null space of G.
- Suppose we want to solve the IFK $d(s) = \int_0^1 g(x,s)\, m(x)\, dx$, given two values $d(s_i)$, $i = 1, 2$, for the parameter $m(x)$. Write out the linear system that results if we use collocation at the points $1/4$ and $3/4$ and the midpoint rule for integration.
- What is the Tikhonov regularization of the (possibly) ill-posed problem $G\mathbf{m} = \mathbf{d}$? Exhibit the Lagrangian of the problem of minimizing $\|G\mathbf{m} - \mathbf{d}\|_2$ subject to the constraint $\|\mathbf{m}\|_2 \leq \epsilon$ and show how it is related to Tikhonov regularization.

Sample Questions:

- Assume that for a full column rank matrix G, $G^T G$ is invertible and $G\mathbf{m}_{true} = \mathbf{d}_{true}$ is satisfied by a model and exact data. Show that if the coordinates of the error vector $\mathbf{d} - \mathbf{d}_{true}$ are independent and identically distributed with mean zero and standard deviation σ, then $E[\mathbf{m}] = \mathbf{m}_{true}$, where $\mathbf{m}$ solves the least squares problem for $G\mathbf{m} = \mathbf{d}$.
- State the SVD Theorem and explain how it enables us to compute the null space of G.
- Suppose we want to solve the IFK $d(s) = \int_0^1 g(x,s)\, m(x)\, dx$, given two values $d(s_i)$, $i = 1, 2$, for the parameter $m(x)$. Write out the linear system that results if we use collocation at the points $1/4$ and $3/4$ and the midpoint rule for integration.
- What is the Tikhonov regularization of the (possibly) ill-posed problem $G\mathbf{m} = \mathbf{d}$? Exhibit the Lagrangian of the problem of minimizing $\|G\mathbf{m} - \mathbf{d}\|_2$ subject to the constraint $\|\mathbf{m}\|_2 \leq \epsilon$ and show how it is related to Tikhonov regularization.

Sample Questions:

- Assume that for a full column rank matrix G, $G^T G$ is invertible and $G\mathbf{m}_{true} = \mathbf{d}_{true}$ is satisfied by a model and exact data. Show that if the coordinates of the error vector $\mathbf{d} - \mathbf{d}_{true}$ are independent and identically distributed with mean zero and standard deviation σ, then $E[\mathbf{m}] = \mathbf{m}_{true}$, where $\mathbf{m}$ solves the least squares problem for $G\mathbf{m} = \mathbf{d}$.
- State the SVD Theorem and explain how it enables us to compute the null space of G.
- Suppose we want to solve the IFK $d(s) = \int_0^1 g(x,s)\, m(x)\, dx$, given two values $d(s_i)$, $i = 1, 2$, for the parameter $m(x)$. Write out the linear system that results if we use collocation at the points $1/4$ and $3/4$ and the midpoint rule for integration.
- What is the Tikhonov regularization of the (possibly) ill-posed problem $G\mathbf{m} = \mathbf{d}$? Exhibit the Lagrangian of the problem of minimizing $\|G\mathbf{m} - \mathbf{d}\|_2$ subject to the constraint $\|\mathbf{m}\|_2 \leq \epsilon$ and show how it is related to Tikhonov regularization.

Sample Questions:

- Assume that for a full column rank matrix G, $G^T G$ is invertible and $G\mathbf{m}_{true} = \mathbf{d}_{true}$ is satisfied by a model and exact data. Show that if the coordinates of the error vector $\mathbf{d} - \mathbf{d}_{true}$ are independent and identically distributed with mean zero and standard deviation σ, then $E[\mathbf{m}] = \mathbf{m}_{true}$, where $\mathbf{m}$ solves the least squares problem for $G\mathbf{m} = \mathbf{d}$.
- State the SVD Theorem and explain how it enables us to compute the null space of G.
- Suppose we want to solve the IFK $d(s) = \int_0^1 g(x,s)\, m(x)\, dx$, given two values $d(s_i)$, $i = 1, 2$, for the parameter $m(x)$. Write out the linear system that results if we use collocation at the points $1/4$ and $3/4$ and the midpoint rule for integration.
- What is the Tikhonov regularization of the (possibly) ill-posed problem $G\mathbf{m} = \mathbf{d}$? Exhibit the Lagrangian of the problem of minimizing $\|G\mathbf{m} - \mathbf{d}\|_2$ subject to the constraint $\|\mathbf{m}\|_2 \leq \epsilon$ and show how it is related to Tikhonov regularization.