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

Instructor: Thomas Shores
Department of Mathematics

AvH 10

Outline

Covariance and Resolution

Definition

The **model resolution matrix** for the problem $G\mathbf{m} = \mathbf{d}$ is $R_{\mathbf{m}} = G^{\dagger}G$.

Consequences:

- $R_{\mathbf{m}} = V_p V_p^T$, which is just I_n if G has full column rank.
- If $G\mathbf{m}_{\mathbf{true}} = \mathbf{d}$, then $E[\mathbf{m}_{\dagger}] = R_m \mathbf{m}_{true}$
- Thus, the bias in the gereralized inverse solution is $E[\mathbf{m}_{\dagger}] - \mathbf{m}_{true} = (R_{\mathbf{m}} - I)\,\mathbf{m}_{true} = -V_0 V_0^T \mathbf{m}_{true}$ with $V = [V_p V_0]$.
- Similarly, in the case of identically distributed data with variance σ^2, the covariance matrix is $\mathrm{Cov}(\mathbf{m}_{\dagger}) = \sigma^2 G^{\dagger}\left(G^{\dagger}\right)^T = \sigma^2 \sum_{i=1}^{p} \frac{\mathbf{v}_i \mathbf{v}_i^T}{\sigma_i^2}$.
- From expected values we obtain a **resolution test**: if a diagonal entry are close to 1, we claim good resolution of that coordinate, otherwise not.

Covariance and Resolution

Definition

The **model resolution matrix** for the problem $G\mathbf{m} = \mathbf{d}$ is $R_{\mathbf{m}} = G^{\dagger}G$.

Consequences:

- $R_{\mathbf{m}} = V_p V_p^T$, which is just I_n if G has full column rank.
- If $G\mathbf{m}_{\mathbf{true}} = \mathbf{d}$, then $E[\mathbf{m}_{\dagger}] = R_m \mathbf{m}_{true}$
- Thus, the bias in the gereralized inverse solution is $E[\mathbf{m}_{\dagger}] - \mathbf{m}_{true} = (R_{\mathbf{m}} - I)\,\mathbf{m}_{true} = -V_0 V_0^T \mathbf{m}_{true}$ with $V = [V_p V_0]$.
- Similarly, in the case of identically distributed data with variance σ^2, the covariance matrix is $\text{Cov}(\mathbf{m}_{\dagger}) = \sigma^2 G^{\dagger}(G^{\dagger})^T = \sigma^2 \sum_{i=1}^{p} \frac{\mathbf{v}_i \mathbf{v}_i^T}{\sigma_i^2}$.
- From expected values we obtain a **resolution test**: if a diagonal entry are close to 1, we claim good resolution of that coordinate, otherwise not.

Covariance and Resolution

Definition

The **model resolution matrix** for the problem $G\mathbf{m} = \mathbf{d}$ is $R_{\mathbf{m}} = G^{\dagger}G$.

Consequences:

- $R_{\mathbf{m}} = V_p V_p^T$, which is just I_n if G has full column rank.
- If $G\mathbf{m}_{\mathbf{true}} = \mathbf{d}$, then $E[\mathbf{m}_{\dagger}] = R_m \mathbf{m}_{true}$
- Thus, the bias in the gereralized inverse solution is $E[\mathbf{m}_{\dagger}] - \mathbf{m}_{true} = (R_{\mathbf{m}} - I)\,\mathbf{m}_{true} = -V_0 V_0^T \mathbf{m}_{true}$ with $V = [V_p V_0]$.
- Similarly, in the case of identically distributed data with variance σ^2, the covariance matrix is $\mathrm{Cov}(\mathbf{m}_{\dagger}) = \sigma^2 G^{\dagger}\left(G^{\dagger}\right)^T = \sigma^2 \sum_{i=1}^{p} \frac{\mathbf{v}_i \mathbf{v}_i^T}{\sigma_i^2}$.
- From expected values we obtain a **resolution test**: if a diagonal entry are close to 1, we claim good resolution of that coordinate, otherwise not.

Covariance and Resolution

Definition

The **model resolution matrix** for the problem $G\mathbf{m} = \mathbf{d}$ is $R_{\mathbf{m}} = G^{\dagger}G$.

Consequences:

- $R_{\mathbf{m}} = V_p V_p^T$, which is just I_n if G has full column rank.
- If $G\mathbf{m_{true}} = \mathbf{d}$, then $E\left[\mathbf{m}_{\dagger}\right] = R_m \mathbf{m}_{true}$
- Thus, the bias in the gereralized inverse solution is $E\left[\mathbf{m}_{\dagger}\right] - \mathbf{m}_{true} = \left(R_{\mathbf{m}} - I\right)\mathbf{m}_{true} = -V_0 V_0^T \mathbf{m}_{true}$ with $V = [V_p V_0]$.
- Similarly, in the case of identically distributed data with variance σ^2, the covariance matrix is $\operatorname{Cov}\left(\mathbf{m}_{\dagger}\right) = \sigma^2 G^{\dagger}\left(G^{\dagger}\right)^T = \sigma^2 \sum_{i=1}^{p} \frac{\mathbf{V}_i \mathbf{V}_i^T}{\sigma_i^2}$.
- From expected values we obtain a **resolution test**: if a diagonal entry are close to 1, we claim good resolution of that coordinate, otherwise not.

Covariance and Resolution

Definition

The **model resolution matrix** for the problem $G\mathbf{m} = \mathbf{d}$ is $R_{\mathbf{m}} = G^{\dagger}G$.

Consequences:

- $R_{\mathbf{m}} = V_p V_p^T$, which is just I_n if G has full column rank.
- If $G\mathbf{m}_{\mathbf{true}} = \mathbf{d}$, then $E[\mathbf{m}_{\dagger}] = R_m \mathbf{m}_{true}$
- Thus, the bias in the gereralized inverse solution is $E[\mathbf{m}_{\dagger}] - \mathbf{m}_{true} = (R_{\mathbf{m}} - I)\,\mathbf{m}_{true} = -V_0 V_0^T \mathbf{m}_{true}$ with $V = [V_p V_0]$.
- Similarly, in the case of identically distributed data with variance σ^2, the covariance matrix is $\mathrm{Cov}(\mathbf{m}_{\dagger}) = \sigma^2 G^{\dagger}\left(G^{\dagger}\right)^T = \sigma^2 \sum_{i=1}^{p} \frac{\mathbf{v}_i \mathbf{v}_i^T}{\sigma_i^2}$.
- From expected values we obtain a **resolution test**: if a diagonal entry are close to 1, we claim good resolution of that coordinate, otherwise not.

Covariance and Resolution

Definition

The **model resolution matrix** for the problem $G\mathbf{m} = \mathbf{d}$ is $R_{\mathbf{m}} = G^{\dagger}G$.

Consequences:

- $R_{\mathbf{m}} = V_p V_p^T$, which is just I_n if G has full column rank.
- If $G\mathbf{m_{true}} = \mathbf{d}$, then $E[\mathbf{m}_{\dagger}] = R_m \mathbf{m}_{true}$
- Thus, the bias in the gereralized inverse solution is $E[\mathbf{m}_{\dagger}] - \mathbf{m}_{true} = (R_{\mathbf{m}} - I)\,\mathbf{m}_{true} = -V_0 V_0^T \mathbf{m}_{true}$ with $V = [V_p V_0]$.
- Similarly, in the case of identically distributed data with variance σ^2, the covariance matrix is $\mathrm{Cov}(\mathbf{m}_{\dagger}) = \sigma^2 G^{\dagger}\left(G^{\dagger}\right)^T = \sigma^2 \sum_{i=1}^{p} \frac{\mathbf{v}_i \mathbf{v}_i^T}{\sigma_i^2}$.
- From expected values we obtain a **resolution test**: if a diagonal entry are close to 1, we claim good resolution of that coordinate, otherwise not.

Outline

Instability of Generalized Inverse Solution

The key results:

- For $n \times n$ square matrix G $\text{cond}_2(G) = \|G\|_2 \left\|G^{-1}\right\|_2 = \sigma_1/\sigma_n$.
- This inspires the definition: the condition number of an $m \times n$ matrix G is σ_1/σ_q where $q = \min\{m, n\}$.
- Note: if $\sigma_q = 0$, the condition number is infinity. Is this notion useful?
- If data $\mathbf{d}$ vector is perturbed to $\mathbf{d}'$, resulting in a perturbation of the generalized inverse solution $\mathbf{m}_\dagger$ to $\mathbf{m}'_\dagger$, then $\frac{\left\|\mathbf{m}'_\dagger - \mathbf{m}_\dagger\right\|_2}{\left\|\mathbf{m}_\dagger\right\|_2} \leq \text{cond}(G) \frac{\|\mathbf{d}' - \mathbf{d}\|_2}{\|\mathbf{d}\|_2}$.

Instability of Generalized Inverse Solution

The key results:

- For $n \times n$ square matrix G $\text{cond}_2(G) = \|G\|_2 \left\|G^{-1}\right\|_2 = \sigma_1/\sigma_n$.
- This inspires the definition: the condition number of an $m \times n$ matrix G is σ_1/σ_q where $q = \min\{m, n\}$.
- Note: if $\sigma_q = 0$, the condition number is infinity. Is this notion useful?
- If data $\mathbf{d}$ vector is perturbed to $\mathbf{d}'$, resulting in a perturbation of the generalized inverse solution $\mathbf{m}_\dagger$ to $\mathbf{m}'_\dagger$, then $\frac{\left\|\mathbf{m}'_\dagger - \mathbf{m}_\dagger\right\|_2}{\left\|\mathbf{m}_\dagger\right\|_2} \leq \text{cond}(G) \frac{\|\mathbf{d}' - \mathbf{d}\|_2}{\|\mathbf{d}\|_2}$.

Instability of Generalized Inverse Solution

The key results:

- For $n \times n$ square matrix G
  $\mathrm{cond}_2(G) = \|G\|_2 \left\|G^{-1}\right\|_2 = \sigma_1/\sigma_n$.
- This inspires the definition: the condition number of an $m \times n$ matrix G is σ_1/σ_q where $q = \min\{m, n\}$.
- Note: if $\sigma_q = 0$, the condition number is infinity. Is this notion useful?
- If data $\mathbf{d}$ vector is perturbed to $\mathbf{d}'$, resulting in a perturbation of the generalized inverse solution $\mathbf{m}_\dagger$ to $\mathbf{m}_\dagger'$, then
  $\frac{\left\|\mathbf{m}_\dagger' - \mathbf{m}_\dagger\right\|_2}{\left\|\mathbf{m}_\dagger\right\|_2} \leq \mathrm{cond}(G) \frac{\|\mathbf{d}' - \mathbf{d}\|_2}{\|\mathbf{d}\|_2}$.

Instability of Generalized Inverse Solution

The key results:

- For $n \times n$ square matrix G $\operatorname{cond}_2(G) = \|G\|_2 \left\|G^{-1}\right\|_2 = \sigma_1/\sigma_n$.
- This inspires the definition: the condition number of an $m \times n$ matrix G is σ_1/σ_q where $q = \min\{m, n\}$.
- Note: if $\sigma_q = 0$, the condition number is infinity. Is this notion useful?
- If data $\mathbf{d}$ vector is perturbed to $\mathbf{d}'$, resulting in a perturbation of the generalized inverse solution $\mathbf{m}_\dagger$ to $\mathbf{m}'_\dagger$, then $\frac{\left\|\mathbf{m}'_\dagger - \mathbf{m}_\dagger\right\|_2}{\left\|\mathbf{m}_\dagger\right\|_2} \leq \operatorname{cond}(G) \frac{\|\mathbf{d}' - \mathbf{d}\|_2}{\|\mathbf{d}\|_2}$.

Instability of Generalized Inverse Solution

The key results:

- For $n \times n$ square matrix G $\text{cond}_2(G) = \|G\|_2 \left\|G^{-1}\right\|_2 = \sigma_1/\sigma_n$.
- This inspires the definition: the condition number of an $m \times n$ matrix G is σ_1/σ_q where $q = \min\{m, n\}$.
- Note: if $\sigma_q = 0$, the condition number is infinity. Is this notion useful?
- If data $\mathbf{d}$ vector is perturbed to $\mathbf{d}'$, resulting in a perturbation of the generalized inverse solution $\mathbf{m}_\dagger$ to $\mathbf{m}'_\dagger$, then $\frac{\left\|\mathbf{m}'_\dagger - \mathbf{m}_\dagger\right\|_2}{\left\|\mathbf{m}_\dagger\right\|_2} \leq \text{cond}(G) \frac{\|\mathbf{d}' - \mathbf{d}\|_2}{\|\mathbf{d}\|_2}$.

Stability Issues

How these facts affect stability:

- If cond(G) is not too large, then the solution is stable to perturbations in data.
- If $\sigma_1 \gg \sigma_p$, there is a potential for instability. It is diminished if the data itself has small components in the direction of singular vectors corresponding to small singular values.
- If $\sigma_1 \gg \sigma_p$, and there is a clear delineation between "small" singular values and the rest, we simple discard the small singular values and treat the problem as one of smaller rank with "good" singular values.
- If $\sigma_1 \gg \sigma_p$, and there is no clear delineation between "small" singular values and the rest, we have to discard some of them, but which ones? This leads to regularization issues. In any case, any method that discards small singular values produces a **truncated SVD** (TSVD) solution.

Stability Issues

How these facts affect stability:

- If cond (G) is not too large, then the solution is stable to perturbations in data.
- If $\sigma_1 \gg \sigma_p$, there is a potential for instability. It is diminished if the data itself has small components in the direction of singular vectors corresponding to small singular values.
- If $\sigma_1 \gg \sigma_p$, and there is a clear delineation between "small" singular values and the rest, we simple discard the small singular values and treat the problem as one of smaller rank with "good" singular values.
- If $\sigma_1 \gg \sigma_p$, and there is no clear delineation between "small" singular values and the rest, we have to discard some of them, but which ones? This leads to regularization issues. In any case, any method that discards small singular values produces a **truncated SVD** (TSVD) solution.

Stability Issues

How these facts affect stability:

- If cond (G) is not too large, then the solution is stable to perturbations in data.
- If $\sigma_1 \gg \sigma_p$, there is a potential for instability. It is diminished if the data itself has small components in the direction of singular vectors corresponding to small singular values.
- If $\sigma_1 \gg \sigma_p$, and there is a clear delineation between "small" singular values and the rest, we simple discard the small singular values and treat the problem as one of smaller rank with "good" singular values.
- If $\sigma_1 \gg \sigma_p$, and there is no clear delineation between "small" singular values and the rest, we have to discard some of them, but which ones? This leads to regularization issues. In any case, any method that discards small singular values produces a **truncated SVD** (TSVD) solution.

Stability Issues

How these facts affect stability:

- If cond (G) is not too large, then the solution is stable to perturbations in data.
- If $\sigma_1 \gg \sigma_p$, there is a potential for instability. It is diminished if the data itself has small components in the direction of singular vectors corresponding to small singular values.
- If $\sigma_1 \gg \sigma_p$, and there is a clear delineation between "small" singular values and the rest, we simple discard the small singular values and treat the problem as one of smaller rank with "good" singular values.
- If $\sigma_1 \gg \sigma_p$, and there is no clear delineation between "small" singular values and the rest, we have to discard some of them, but which ones? This leads to regularization issues. In any case, any method that discards small singular values produces a **truncated SVD** (TSVD) solution.

Stability Issues

How these facts affect stability:

- If cond (G) is not too large, then the solution is stable to perturbations in data.
- If $\sigma_1 \gg \sigma_p$, there is a potential for instability. It is diminished if the data itself has small components in the direction of singular vectors corresponding to small singular values.
- If $\sigma_1 \gg \sigma_p$, and there is a clear delineation between "small" singular values and the rest, we simple discard the small singular values and treat the problem as one of smaller rank with "good" singular values.
- If $\sigma_1 \gg \sigma_p$, and there is no clear delineation between "small" singular values and the rest, we have to discard some of them, but which ones? This leads to regularization issues. In any case, any method that discards small singular values produces a **truncated SVD** (TSVD) solution.

Outline

1. Chapter 4: Rank Deficiency and Ill-Conditioning
   - Covariance and Resolution of the Generalized Inverse Solution
   - Instability of Generalized Inverse Solutions
   - An Example of a Rank-Deficient Problem
   - Discrete Ill-Posed Problems

Linear Tomography Models

Note: Rank deficient problems are *automatically* ill-posed.

Basic Idea:

A ray emanates from one known point to another along a known path ℓ, with a detectable property which is observable data. These data are used to estimate a travel property of the medium. For example, let the property be travel time, so that:

- Travel time is given by $t = \int_{\ell} \frac{dt}{dx} dx = \int_{\ell} \frac{1}{v(x)} dx$
- We can linearize by making paths straight lines.
- Discretize by embedding the medium in a square (cube) and subdividing it into regular subsquares (cubes) in which we assume “slowness” (parameter of the problem) is constant.
- Transmit the ray along specified paths and collect temporal data to be used in estimating “slowness”.

Linear Tomography Models

Note: Rank deficient problems are *automatically* ill-posed.

Basic Idea:

A ray emanates from one known point to another along a known path ℓ, with a detectable property which is observable data. These data are used to estimate a travel property of the medium. For example, let the property be travel time, so that:

- Travel time is given by $t = \int_\ell \frac{dt}{dx} dx = \int_\ell \frac{1}{v(x)} dx$
- We can linearize by making paths straight lines.
- Discretize by embedding the medium in a square (cube) and subdividing it into regular subsquares (cubes) in which we assume "slowness" (parameter of the problem) is constant.
- Transmit the ray along specified paths and collect temporal data to be used in estimating "slowness".

Linear Tomography Models

Note: Rank deficient problems are *automatically* ill-posed.

Basic Idea:

A ray emanates from one known point to another along a known path ℓ, with a detectable property which is observable data. These data are used to estimate a travel property of the medium. For example, let the property be travel time, so that:

- Travel time is given by $t = \int_\ell \frac{dt}{dx} dx = \int_\ell \frac{1}{v(x)} dx$
- We can linearize by making paths straight lines.
- Discretize by embedding the medium in a square (cube) and subdividing it into regular subsquares (cubes) in which we assume "slowness" (parameter of the problem) is constant.
- Transmit the ray along specified paths and collect temporal data to be used in estimating "slowness".

Linear Tomography Models

Note: Rank deficient problems are *automatically* ill-posed.

Basic Idea:

A ray emanates from one known point to another along a known path ℓ, with a detectable property which is observable data. These data are used to estimate a travel property of the medium. For example, let the property be travel time, so that:

- Travel time is given by $t = \int_\ell \frac{dt}{dx} dx = \int_\ell \frac{1}{v(x)} dx$
- We can linearize by making paths straight lines.
- Discretize by embedding the medium in a square (cube) and subdividing it into regular subsquares (cubes) in which we assume "slowness" (parameter of the problem) is constant.
- Transmit the ray along specified paths and collect temporal data to be used in estimating "slowness".

Linear Tomography Models

Note: Rank deficient problems are *automatically* ill-posed.

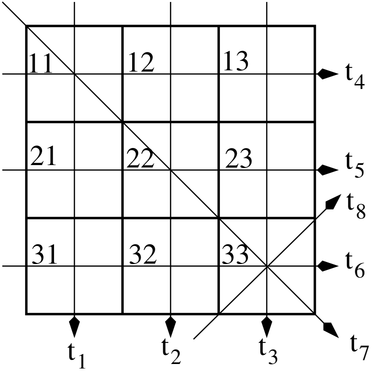
Basic Idea:

A ray emanates from one known point to another along a known path ℓ, with a detectable property which is observable data. These data are used to estimate a travel property of the medium. For example, let the property be travel time, so that:

- Travel time is given by $t = \int_\ell \frac{dt}{dx} dx = \int_\ell \frac{1}{v(x)} dx$
- We can linearize by making paths straight lines.
- Discretize by embedding the medium in a square (cube) and subdividing it into regular subsquares (cubes) in which we assume "slowness" (parameter of the problem) is constant.
- Transmit the ray along specified paths and collect temporal data to be used in estimating "slowness".

Example 1.6 and 4.1

The figure for this experiment (assume each subsquare has sides of length 1, so the size of the large square is 3×3):



Example 1.6 and 4.1

Corresponding matrix of distances G (rows of G represent distances along corresponding path, columns the ray distances across each subblock) and resulting system:

$$G\mathbf{m} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ \sqrt{2} & 0 & 0 & 0 & \sqrt{2} & 0 & 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} s_{11} \\ s_{12} \\ s_{13} \\ s_{21} \\ s_{22} \\ s_{23} \\ s_{31} \\ s_{32} \\ s_{33} \end{bmatrix} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \\ t_7 \\ t_8 \end{bmatrix} = \mathbf{d}$$

Observe: in this Example $m = 8$ and $n = 9$. So it is clearly rank deficient. Now let's run the example file for this example. View and discuss the source.

Outline

What Are They?

These problems arise due to ill-conditioning of G, *as opposed to a rank deficiency problem. Theoretically*, they are not ill-posed, like the Hilbert matrix. But practically speaking, they behave like ill-posed problems. Authors present a hierarchy of sorts for a problem with system $G\mathbf{m} = \mathbf{d}$. These order expressions are valid as $j \to \infty$.

- $\mathcal{O}\left(\frac{1}{j^\alpha}\right)$ with $0 < \alpha \leq 1$, the problem is **mildly** ill-posed.
- $\mathcal{O}\left(\frac{1}{j^\alpha}\right)$ with $\alpha > 1$, the problem is **moderately** ill-posed.
- $\mathcal{O}\left(e^{-\alpha j}\right)$ with $0 < \alpha$, the problem is **severely** ill-posed.

What Are They?

These problems arise due to ill-conditioning of G, *as opposed to a rank deficiency problem. Theoretically*, they are not ill-posed, like the Hilbert matrix. But practically speaking, they behave like ill-posed problems. Authors present a hierarchy of sorts for a problem with system $G\mathbf{m} = \mathbf{d}$. These order expressions are valid as $j \to \infty$.

- $\mathcal{O}\left(\frac{1}{j^\alpha}\right)$ with $0 < \alpha \le 1$, the problem is **mildly** ill-posed.
- $\mathcal{O}\left(\frac{1}{j^\alpha}\right)$ with $\alpha > 1$, the problem is **moderately** ill-posed.
- $\mathcal{O}\left(e^{-\alpha j}\right)$ with $0 < \alpha$, the problem is **severely** ill-posed.

What Are They?

These problems arise due to ill-conditioning of G, *as opposed to a rank deficiency problem.* *Theoretically*, they are not ill-posed, like the Hilbert matrix. But practically speaking, they behave like ill-posed problems. Authors present a hierarchy of sorts for a problem with system $G\mathbf{m} = \mathbf{d}$. These order expressions are valid as $j \to \infty$.

- $\mathcal{O}\left(\frac{1}{j^\alpha}\right)$ with $0 < \alpha \leq 1$, the problem is **mildly** ill-posed.
- $\mathcal{O}\left(\frac{1}{j^\alpha}\right)$ with $\alpha > 1$, the problem is **moderately** ill-posed.
- $\mathcal{O}\left(e^{-\alpha j}\right)$ with $0 < \alpha$, the problem is **severely** ill-posed.

What Are They?

These problems arise due to ill-conditioning of G, *as opposed to a rank deficiency problem. Theoretically*, they are not ill-posed, like the Hilbert matrix. But practically speaking, they behave like ill-posed problems. Authors present a hierarchy of sorts for a problem with system $G\mathbf{m} = \mathbf{d}$. These order expressions are valid as $j \to \infty$.

- $\mathcal{O}\left(\frac{1}{j^\alpha}\right)$ with $0 < \alpha \leq 1$, the problem is **mildly** ill-posed.
- $\mathcal{O}\left(\frac{1}{j^\alpha}\right)$ with $\alpha > 1$, the problem is **moderately** ill-posed.
- $\mathcal{O}\left(e^{-\alpha j}\right)$ with $0 < \alpha$, the problem is **severely** ill-posed.

A Severly Ill-Posed Problem

The Shaw Problem:

An optics experiment is performed by dividing a circle using a vertical transversal with a slit in the middle. A variable intensity light source is placed around the left half of the circle and rays pass through the slit, where they are measured at points on the right half of the circle.

- Measure angles counterclockwise from the x-axis, using $-\pi/2 \leq \theta \leq \pi/2$ for the source intensity $m(\theta)$, and $-\pi/2 \leq s \leq \pi/2$ for destination intensity $d(s)$.
- The model for this problem comes from diffraction theory: $d(s) = \int_{-\pi/2}^{\pi/2} (\cos(s) + \cos(\theta))^2 \left(\frac{\sin(\pi(\sin(s) + \sin(\theta)))}{\pi(\sin(s) + \sin(\theta))} \right)^2 m(\theta)\, d\theta.$

A Severly Ill-Posed Problem

The Shaw Problem:

An optics experiment is performed by dividing a circle using a vertical transversal with a slit in the middle. A variable intensity light source is placed around the left half of the circle and rays pass through the slit, where they are measured at points on the right half of the circle.

- Measure angles counterclockwise from the x-axis, using $-\pi/2 \leq \theta \leq \pi/2$ for the source intensity $m(\theta)$, and $-\pi/2 \leq s \leq \pi/2$ for destination intensity $d(s)$.
- The model for this problem comes from diffraction theory: $d(s) =$
$$\int_{-\pi/2}^{\pi/2} (\cos(s) + \cos(\theta))^2 \left(\frac{\sin(\pi(\sin(s) + \sin(\theta)))}{\pi(\sin(s) + \sin(\theta))} \right)^2 m(\theta)\, d\theta.$$

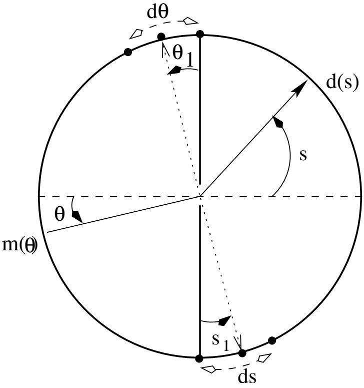
A Severly Ill-Posed Problem

The Shaw Problem:

An optics experiment is performed by dividing a circle using a vertical transversal with a slit in the middle. A variable intensity light source is placed around the left half of the circle and rays pass through the slit, where they are measured at points on the right half of the circle.

- Measure angles counterclockwise from the x-axis, using $-\pi/2 \leq \theta \leq \pi/2$ for the source intensity $m(\theta)$, and $-\pi/2 \leq s \leq \pi/2$ for destination intensity $d(s)$.
- The model for this problem comes from diffraction theory:
$$d(s) = \int_{-\pi/2}^{\pi/2} (\cos(s) + \cos(\theta))^2 \left(\frac{\sin(\pi(\sin(s) + \sin(\theta)))}{\pi(\sin(s) + \sin(\theta))} \right)^2 m(\theta)\, d\theta.$$

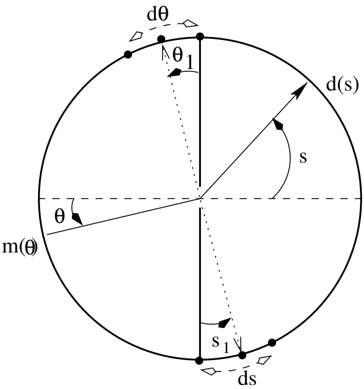
The Shaw Problem



Two Problems:

- The forward problem: given source intensity $m(\theta)$, compute the destination intensity $d(s)$.
- The inverse problem: given destination intensity $d(s)$, compute the source intensity $m(\theta)$.
- It can be shown that the inverse problem is severly ill-posed.

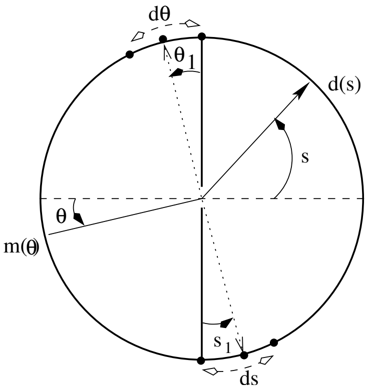
The Shaw Problem



Two Problems:

- The forward problem: given source intensity $m(\theta)$, compute the destination intensity $d(s)$.
- The inverse problem: given destination intensity $d(s)$, compute the source intensity $m(\theta)$.
- It can be shown that the inverse problem is severly ill-posed.

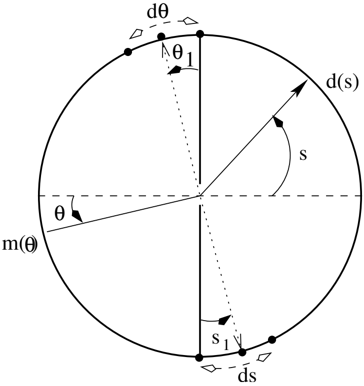
The Shaw Problem



Two Problems:

- The forward problem: given source intensity $m(\theta)$, compute the destination intensity $d(s)$.
- The inverse problem: given destination intensity $d(s)$, compute the source intensity $m(\theta)$.
- It can be shown that the inverse problem is severely ill-posed.

The Shaw Problem



Two Problems:

- The forward problem: given source intensity $m(\theta)$, compute the destination intensity $d(s)$.
- The inverse problem: given destination intensity $d(s)$, compute the source intensity $m(\theta)$.
- It can be shown that the inverse problem is severly ill-posed.

How To Discretize The Problem:

- Discretize the parameter domain $-\pi/2 \leq \theta \leq \pi/2$ and the data domain $-\pi/2 \leq s \leq \pi/2$ into n subintervals of equal size $\Delta s = \Delta\theta = \pi/n$.
- Therefore, and let s_i, θ_i be the midpoints of the i-th subintervals:

$$s_i = \theta_i = -\frac{\pi}{2} + \frac{(i - 0.5)\,\pi}{2}, \; i = 1, 2, \ldots, n.$$

- Define

$$G_{i,j} = (\cos(s_i) + \cos(\theta_i))^2 \left(\frac{\sin(\pi(\sin(s_i) + \sin(\theta_i)))}{\pi(\sin(s_i) + \sin(\theta_i))} \right)^2 \Delta\theta$$

- Thus if $m_i \approx m(\theta_i)$, $d_i \approx d(s_i)$, $\mathbf{m} = (m_1, m_2, \ldots, m_n)$ and $\mathbf{d} = (d_1, d_2, \ldots, d_n)$, then discretization and the midpoint rule give $G\mathbf{m} = \mathbf{d}$, as in Chapter 3.

The Shaw Problem

How To Discretize The Problem:

- Discretize the parameter domain $-\pi/2 \leq \theta \leq \pi/2$ and the data domain $-\pi/2 \leq s \leq \pi/2$ into n subintervals of equal size $\Delta s = \Delta\theta = \pi/n$.
- Therefore, and let s_i, θ_i be the midpoints of the i-th subintervals:

$$s_i = \theta_i = -\frac{\pi}{2} + \frac{(i-0.5)\,\pi}{2}, \; i = 1, 2, \ldots, n.$$

- Define

$$G_{i,j} = (\cos(s_i) + \cos(\theta_i))^2 \left(\frac{\sin(\pi(\sin(s_i) + \sin(\theta_i)))}{\pi(\sin(s_i) + \sin(\theta_i))} \right)^2 \Delta\theta$$

- Thus if $m_i \approx m(\theta_i)$, $d_i \approx d(s_i)$, $\mathbf{m} = (m_1, m_2, \ldots, m_n)$ and $\mathbf{d} = (d_1, d_2, \ldots, d_n)$, then discretization and the midpoint rule give $G\mathbf{m} = \mathbf{d}$, as in Chapter 3.

How To Discretize The Problem:

- Discretize the parameter domain $-\pi/2 \leq \theta \leq \pi/2$ and the data domain $-\pi/2 \leq s \leq \pi/2$ into n subintervals of equal size $\Delta s = \Delta\theta = \pi/n$.
- Therefore, and let s_i, θ_i be the midpoints of the i-th subintervals:

$$s_i = \theta_i = -\frac{\pi}{2} + \frac{(i - 0.5)\,\pi}{2},\ i = 1, 2, \ldots, n.$$

- Define

$$G_{i,j} = (\cos(s_i) + \cos(\theta_i))^2 \left(\frac{\sin(\pi(\sin(s_i) + \sin(\theta_i)))}{\pi(\sin(s_i) + \sin(\theta_i))}\right)^2 \Delta\theta$$

- Thus if $m_i \approx m(\theta_i)$, $d_i \approx d(s_i)$, $\mathbf{m} = (m_1, m_2, \ldots, m_n)$ and $\mathbf{d} = (d_1, d_2, \ldots, d_n)$, then discretization and the midpoint rule give $G\mathbf{m} = \mathbf{d}$, as in Chapter 3.

How To Discretize The Problem:

- Discretize the parameter domain $-\pi/2 \le \theta \le \pi/2$ and the data domain $-\pi/2 \le s \le \pi/2$ into n subintervals of equal size $\Delta s = \Delta\theta = \pi/n$.
- Therefore, and let s_i, θ_i be the midpoints of the i-th subintervals:

$$s_i = \theta_i = -\frac{\pi}{2} + \frac{(i - 0.5)\,\pi}{2}, \; i = 1, 2, \ldots, n.$$

- Define

$$G_{i,j} = (\cos(s_i) + \cos(\theta_i))^2 \left(\frac{\sin(\pi(\sin(s_i) + \sin(\theta_i)))}{\pi(\sin(s_i) + \sin(\theta_i))} \right)^2 \Delta\theta$$

- Thus if $m_i \approx m(\theta_i)$, $d_i \approx d(s_i)$, $\mathbf{m} = (m_1, m_2, \ldots, m_n)$ and $\mathbf{d} = (d_1, d_2, \ldots, d_n)$, then discretization and the midpoint rule give $G\mathbf{m} = \mathbf{d}$, as in Chapter 3.

The Shaw Problem

How To Discretize The Problem:

- Discretize the parameter domain $-\pi/2 \leq \theta \leq \pi/2$ and the data domain $-\pi/2 \leq s \leq \pi/2$ into n subintervals of equal size $\Delta s = \Delta\theta = \pi/n$.
- Therefore, and let s_i, θ_i be the midpoints of the i-th subintervals:

$$s_i = \theta_i = -\frac{\pi}{2} + \frac{(i - 0.5)\,\pi}{2}, \; i = 1, 2, \ldots, n.$$

- Define

$$G_{i,j} = (\cos(s_i) + \cos(\theta_i))^2 \left(\frac{\sin(\pi(\sin(s_i) + \sin(\theta_i)))}{\pi(\sin(s_i) + \sin(\theta_i))} \right)^2 \Delta\theta$$

- Thus if $m_i \approx m(\theta_i)$, $d_i \approx d(s_i)$, $\mathbf{m} = (m_1, m_2, \ldots, m_n)$ and $\mathbf{d} = (d_1, d_2, \ldots, d_n)$, then discretization and the midpoint rule give $G\mathbf{m} = \mathbf{d}$, as in Chapter 3.

The Shaw Problem

Now we can examine the example files on the text CD for this problem. However, here's an easy way to build the matrix G without loops. Basically, these tools were designed to help with 3-D plotting.

```
> n = 20
> ds = pi/n
> s = linspace(ds, pi - ds,n)
> theta = s;
> [S, Theta] = meshgrid(s,theta);
>G = (cos(S) + cos(Theta)).^2 .* (sin(pi*(sin(S) + ...
sin(Theta)))./(pi*(sin(S) + sin(Theta)))).^2*ds;
> % want to see G(s,θ)?
> mesh(S,Theta,G)
> cond(G)
> svd(G)
> rank(G)
```