Hybrid LSQR and RSVD solutions of Ill-Posed Least Squares Problems

Rosemary Renaut\textsuperscript{1} Anthony Helmstetter\textsuperscript{1} Saeed Vatankhah\textsuperscript{2}

\textsuperscript{1}: School of Mathematical and Statistical Sciences, Arizona State University, renaut@asu.edu, anthony.helmstetter@asu.edu

\textsuperscript{2}: Institute of Geophysics, University of Tehran, svatan@ut.ac.ir

Syracuse University
Aims

Motivation Example: Large Scale Gravity Inversion

Background: SVD for the small scale
Standard Approaches to Estimate Regularization Problem

Methods for the Large Scale: Approximating the SVD
Krylov: Golub Kahan Bidiagonalization - LSQR
Randomized SVD

Simulations
1D Contrasting RSVD and LSQR
Two dimensional Examples

Extension: Iteratively Reweighted Regularization [LK83]
Undersampled 3D Magnetic: approximate $L1$ regularization

Conclusions: RSVD - LSQR
Aims

Motivation Example: Large Scale Gravity Inversion

Background: SVD for the small scale

Standard Approaches to Estimate Regularization Problem

Methods for the Large Scale: Approximating the SVD

Krylov: Golub Kahan Bidiagonalization - LSQR

Randomized SVD

Simulations

$1D$ Contrasting RSVD and LSQR

Two dimensional Examples

Extension: Iteratively Reweighted Regularization [LK83]

Undersampled $3D$ Magnetic: approximate $L_1$ regularization

Conclusions: RSVD - LSQR
Aims
Motivation Example: Large Scale Gravity Inversion

Background: SVD for the small scale
Standard Approaches to Estimate Regularization Problem

Methods for the Large Scale: Approximating the SVD
Krylov: Golub Kahan Bidiagonalization - LSQR
Randomized SVD
Simulations
  1D Contrasting RSVD and LSQR
  Two dimensional Examples
Extension: Iteratively Reweighted Regularization [LK83]
  Undersampled 3D Magnetic: approximate $L_1$ regularization

Conclusions: RSVD - LSQR
Outline

Aims
  Motivation Example: Large Scale Gravity Inversion

Background: SVD for the small scale
  Standard Approaches to Estimate Regularization Problem

Methods for the Large Scale: Approximating the SVD
  Krylov: Golub Kahan Bidiagonalization - LSQR
  Randomized SVD

Simulations
  $1D$ Contrasting RSVD and LSQR
  Two dimensional Examples

Extension: Iteratively Reweighted Regularization [LK83]
  Undersampled $3D$ Magnetic: approximate $L1$ regularization

Conclusions: RSVD - LSQR
Research Aims: Efficient solvers for 3D inversion

1. Effective regularization parameter estimation using small scale surrogate for large scale problem
2. Derive surrogates from Krylov projection - LSQR
3. Derive surrogates from Randomized Singular Value Decomposition (RSVD)
4. Effective implementation of $L_1$ and $L_p$ regularization using surrogate initialization.

Explain: hopefully - what these statements mean
Research Aims: Efficient solvers for 3D inversion

1. Effective regularization parameter estimation using small scale surrogate for large scale problem
2. Derive surrogates from Krylov projection - LSQR
3. Derive surrogates from Randomized Singular Value Decomposition (RSVD)
4. Effective implementation of $L_1$ and $L_p$ regularization using surrogate initialization.

Explain: hopefully - what these statements mean
Research Aims: Efficient solvers for 3D inversion

1. Effective regularization parameter estimation using small scale surrogate for large scale problem
2. Derive surrogates from Krylov projection - LSQR
3. Derive surrogates from Randomized Singular Value Decomposition (RSVD)
4. Effective implementation of $L_1$ and $L_p$ regularization using surrogate initialization.

Explain: hopefully - what these statements mean
Research Aims: Efficient solvers for 3D inversion

1. Effective regularization parameter estimation using small scale surrogate for large scale problem
2. Derive surrogates from Krylov projection - LSQR
3. Derive surrogates from Randomized Singular Value Decomposition (RSVD)
4. Effective implementation of $L_1$ and $L_p$ regularization using surrogate initialization.

Explain: hopefully - what these statements mean
Research Aims: Efficient solvers for 3D inversion

1. Effective regularization parameter estimation using small scale surrogate for large scale problem
2. Derive surrogates from Krylov projection - LSQR
3. Derive surrogates from Randomized Singular Value Decomposition (RSVD)
4. Effective implementation of $L_1$ and $L_p$ regularization using surrogate initialization.

Explain: hopefully - what these statements mean
Observation point \( \mathbf{r} = (x, y, z) \)

Vertical magnetic anomaly \( m(\mathbf{r}) \)

\[
m(\mathbf{r}) \propto \int_{d\Omega} \kappa(\mathbf{r}') \frac{\mathbf{r}' - \mathbf{r}}{|\mathbf{r}' - \mathbf{r}|^3} d\mathbf{r}'
\]

Susceptibility \( \kappa(\mathbf{r}') \) at \( \mathbf{r}' = (x', y', z') \)

Linear Relation \( \mathbf{m} = G\kappa \)

Aim: Given surface observations \( m_{ij} \) find susceptibility \( \kappa_{ijk} \)

Underdetermined, measurements 5000, unknowns 75000

Practical Approaches for Large Scale Ill-Posed Problems needed
Motivation Example: Large Scale 3D Magnetic Anomaly Inversion

Observation point \( \mathbf{r} = (x, y, z) \)

Vertical magnetic anomaly \( m(\mathbf{r}) \)

\[
m(\mathbf{r}) \propto \int_{d\Omega} \kappa(\mathbf{r}') \frac{\mathbf{r}' - \mathbf{r}}{|\mathbf{r}' - \mathbf{r}|^3} \, d\Omega'
\]

Susceptibility \( \kappa(\mathbf{r}') \) at \( \mathbf{r}' = (x', y', z') \)

Linear Relation \( \mathbf{m} = G\kappa \)

Aim: Given surface observations \( m_{ij} \) find susceptibility \( \kappa_{ijk} \)

Underdetermined, measurements 5000, unknowns 75000

Practical Approaches for Large Scale Ill-Posed Problems needed
Motivation Example: Large Scale 3D Magnetic Anomaly Inversion

Observation point  \( \mathbf{r} = (x, y, z) \)

Vertical magnetic anomaly  \( m(\mathbf{r}) \)

\[
m(\mathbf{r}) \propto \int d\Omega \kappa(\mathbf{r}') \frac{\mathbf{r}' - \mathbf{r}}{|\mathbf{r}' - \mathbf{r}|^3} d' \]

Susceptibility  \( \kappa(\mathbf{r'}) \) at  \( \mathbf{r}' = (x', y', z') \)

Linear Relation  \( \mathbf{m} = G\kappa \)

Aim: Given surface observations  \( m_{ij} \) find susceptibility  \( \kappa_{ijk} \)

Underdetermined, measurements  \( 5000 \), unknowns  \( 75000 \)

Practical Approaches for Large Scale Ill-Posed Problems needed
Motivation Example: Large Scale $3D$ Magnetic Anomaly Inversion

Observation point $r = (x, y, z)$

Vertical magnetic anomaly $m(r)$

$$m(r) \propto \int_{d\Omega} \kappa(r') \frac{r' - r}{|r' - r|^3} d\Omega'$$

Susceptibility $\kappa(r')$ at $r' = (x', y', z')$

Linear Relation $m = G\kappa$

Aim: Given surface observations $m_{ij}$ find susceptibility $\kappa_{ijk}$

Underdetermined, measurements 5000, unknowns 75000

Practical Approaches for Large Scale Ill-Posed Problems needed
Motivation Example: Large Scale 3D Magnetic Anomaly Inversion

Observation point \( \mathbf{r} = (x, y, z) \)

Vertical magnetic anomaly \( m(\mathbf{r}) \)

\[
m(\mathbf{r}) \propto \int d\Omega \kappa(\mathbf{r}') \frac{\mathbf{r}' - \mathbf{r}}{|\mathbf{r}' - \mathbf{r}|^3} d\Omega'
\]

Susceptibility \( \kappa(\mathbf{r}') \) at \( \mathbf{r}' = (x', y', z') \)

Linear Relation \( \mathbf{m} = G\kappa \)

**Aim:** Given surface observations \( m_{ij} \) find susceptibility \( \kappa_{ijk} \)

**Underdetermined, measurements** 5000, **unknowns** 75000

Practical Approaches for Large Scale Ill-Posed Problems needed
Motivation Example: Large Scale $3D$ Magnetic Anomaly Inversion

Observation point $\mathbf{r} = (x, y, z)$

Vertical magnetic anomaly $m(\mathbf{r})$

\[
m(\mathbf{r}) \propto \int d\Omega \kappa(\mathbf{r}') \frac{\mathbf{r}' - \mathbf{r}}{|\mathbf{r}' - \mathbf{r}|^3} d\Omega'
\]

Susceptibility $\kappa(\mathbf{r}')$ at $\mathbf{r}' = (x', y', z')$

Linear Relation $\mathbf{m} = G\kappa$

Aim: Given surface observations $m_{ij}$ find susceptibility $\kappa_{ijk}$

Underdetermined, measurements $5000$, unknowns $75000$

Practical Approaches for Large Scale Ill-Posed Problems needed
Ill-Posed Problem: Example Solutions Image Restoration

True

Contaminated

Naive Restoration
Consider general discrete problem

\[ Ax = b, \quad A \in \mathbb{R}^{m \times n}, \quad b \in \mathbb{R}^m, \quad x \in \mathbb{R}^n. \]

Singular value decomposition (SVD) of \( A \) rank \( r \leq \min(m, n) \)

\[
A = U \Sigma V^T = \sum_{i=1}^{r} u_i \sigma_i v_i^T, \quad \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_r).
\]

Singular values \( \sigma_i = \sqrt{\lambda_i}, \lambda_i \) the eigenvalue of \( A^T A \).
Singular Vectors \( u_i, v_i \):

\[
R(A) = \text{span}(U(:, 1 : r)) \quad N(A) = \text{span}(V(:, r + 1 : n))
\]

Expansion for the solution

\[
x = \sum_{i=1}^{r} \frac{u_i^T b}{\sigma_i} v_i
\]
Background Spectral Decomposition of the Solution: The SVD

Consider general discrete problem

\[ A \mathbf{x} = \mathbf{b}, \quad A \in \mathbb{R}^{m \times n}, \quad \mathbf{b} \in \mathbb{R}^{m}, \quad \mathbf{x} \in \mathbb{R}^{n}. \]

Singular value decomposition (SVD) of \( A \) rank \( r \leq \min(m, n) \)

\[ A = U \Sigma V^{T} = \sum_{i=1}^{r} u_{i} \sigma_{i} v_{i}^{T}, \quad \Sigma = \text{diag}(\sigma_{1}, \ldots, \sigma_{r}). \]

Singular values \( \sigma_{i} = \sqrt{\lambda_{i}} \), \( \lambda_{i} \) the eigenvalue of \( A^{T}A \).
Singular Vectors \( u_{i}, v_{i} \):
\( R(A) = \text{span}(U(:, 1 : r)) \) \( N(A) = \text{span}(V(:, r + 1 : n)) \)

Expansion for the solution

\[ \mathbf{x} = \sum_{i=1}^{r} \frac{u_{i}^{T} \mathbf{b}}{\sigma_{i}} v_{i}. \]
Consider general discrete problem

\[ Ax = b, \quad A \in \mathbb{R}^{m \times n}, \quad b \in \mathbb{R}^m, \quad x \in \mathbb{R}^n. \]

Singular value decomposition (SVD) of \( A \) rank \( r \leq \min(m, n) \)

\[ A = U \Sigma V^T = \sum_{i=1}^{r} u_i \sigma_i v_i^T, \quad \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_r). \]

Singular values \( \sigma_i = \sqrt{\lambda_i} \), \( \lambda_i \) the eigenvalue of \( A^T A \).

Singular Vectors \( u_i, v_i \):

\[ R(A) = \text{span}(U(:,1:r)) \quad N(A) = \text{span}(V(:,r+1:n)) \]

Expansion for the solution

\[ x = \sum_{i=1}^{r} \frac{u_i^T b}{\sigma_i} v_i \]
Consider general discrete problem

\[ A\mathbf{x} = \mathbf{b}, \quad A \in \mathbb{R}^{m \times n}, \quad \mathbf{b} \in \mathbb{R}^{m}, \quad \mathbf{x} \in \mathbb{R}^{n}. \]

Singular value decomposition (SVD) of \( A \) rank \( r \leq \min(m, n) \)

\[ A = U\Sigma V^T = \sum_{i=1}^{r} u_i \sigma_i v_i^T, \quad \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_r). \]

Singular values \( \sigma_i = \sqrt{\lambda_i} \), \( \lambda_i \) the eigenvalue of \( A^T A \).

Singular Vectors \( u_i, v_i \):

\( \text{R}(A) = \text{span}(U(:, 1 : r)) \) \quad \text{N}(A) = \text{span}(V(:, r + 1 : n)) \)

Expansion for the solution

\[ \mathbf{x} = \sum_{i=1}^{r} \frac{u_i^T \mathbf{b}}{\sigma_i} v_i \]
Consider general discrete problem

\[ Ax = b, \quad A \in \mathbb{R}^{m \times n}, \quad b \in \mathbb{R}^m, \quad x \in \mathbb{R}^n. \]

**Singular value decomposition (SVD) of \( A \) rank \( r \leq \min(m, n) \)**

\[ A = U \Sigma V^T = \sum_{i=1}^{r} u_i \sigma_i v_i^T, \quad \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_r). \]

Singular values \( \sigma_i = \sqrt{\lambda_i} \), \( \lambda_i \) the eigenvalue of \( A^T A \).

**Singular Vectors** \( u_i, v_i \):

\[ R(A) = \text{span}(U(:, 1 : r)) \quad N(A) = \text{span}(V(:, r + 1 : n)) \]

Expansion for the solution

\[ x = \sum_{i=1}^{r} \frac{u_i^T b}{\sigma_i} v_i. \]
The Picard Plot - examines the solution

\[ x = \sum_{i=1}^{r} \frac{u_i^T b}{\sigma_i} v_i \]

Solution with these coefficients would be non feasible
The Picard Plot - examines the solution

\[ x = \sum_{i=1}^{r} \frac{u_i^T b}{\sigma_i} v_i \]

Solution with these coefficients would be non feasible
Regularization: Solutions are not stable

**Truncation:** $k < r$ - surrogate problem is size $k$, $A_k \approx A$

\[
x = \sum_{i=1}^{k} \frac{u_i^T b}{\sigma_i} v_i
\]

**Filtering:** $\gamma_i$

\[
x = \sum_{i=1}^{r} \gamma_i(\alpha) \frac{u_i^T b}{\sigma_i} v_i
\]

**Filtered Truncated:** $k, \gamma_i$ - surrogate problem is size $k$, $A_k \approx A$

\[
x = \sum_{i=1}^{k} \gamma_i(\alpha) \frac{u_i^T b}{\sigma_i} v_i
\]

$k$ and $\alpha$ are regularization parameters. $\gamma_i$ filter function
Regularization: Solutions are not stable

**Truncation:** \( k < r \) - surrogate problem is size \( k \), \( A_k \approx A \)

\[
x = \sum_{i=1}^{k} \frac{u_i^T b}{\sigma_i} v_i
\]

**Filtering:** \( \gamma_i \)

\[
x = \sum_{i=1}^{r} \gamma_i(\alpha) \frac{u_i^T b}{\sigma_i} v_i
\]

**Filtered Truncated:** \( k, \gamma_i \) - surrogate problem is size \( k \), \( A_k \approx A \)

\[
x = \sum_{i=1}^{k} \gamma_i(\alpha) \frac{u_i^T b}{\sigma_i} v_i
\]

\( k \) and \( \alpha \) are regularization parameters. \( \gamma_i \) filter function
Regularization: Solutions are not stable

**Truncation:** $k < r$ - surrogate problem is size $k$, $A_k \approx A$

\[
x = \sum_{i=1}^{k} \frac{u_i^T b}{\sigma_i} v_i
\]

**Filtering:** $\gamma_i$

\[
x = \sum_{i=1}^{r} \gamma_i(\alpha) \frac{u_i^T b}{\sigma_i} v_i
\]

**Filtered Truncated:** $k, \gamma_i$ - surrogate problem is size $k$, $A_k \approx A$

\[
x = \sum_{i=1}^{k} \gamma_i(\alpha) \frac{u_i^T b}{\sigma_i} v_i
\]

$k$ and $\alpha$ are regularization parameters. $\gamma_i$ filter function
Regularization: Solutions are not stable

**Truncation:** $k < r$ - surrogate problem is size $k$, $A_k \approx A$

$$x = \sum_{i=1}^{k} \left[ \frac{u_i^T b}{\sigma_i} \right] \nu_i$$

**Filtering:** $\gamma_i$

$$x = \sum_{i=1}^{r} \gamma_i(\alpha) \left[ \frac{u_i^T b}{\sigma_i} \right] \nu_i$$

**Filtered Truncated:** $k, \gamma_i$ - surrogate problem is size $k$, $A_k \approx A$

$$x = \sum_{i=1}^{k} \gamma_i(\alpha) \left[ \frac{u_i^T b}{\sigma_i} \right] \nu_i$$

$k$ and $\alpha$ are regularization parameters. $\gamma_i$ filter function
Tikhonov Regularization: regularization parameter $\alpha$

Filter Functions

$$\gamma_i(\alpha) = \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2}, \ i = 1 \ldots r,$$

Solves Standard Form

$$x(\alpha) = \arg\min_x \{\|b - Ax\|^2 + \alpha^2 \|x\|^2\}$$

$$\approx \arg\min_x \{\|b - A_k x\|^2 + \alpha^2 \|x\|^2\}$$

Generalized Tikhonov has operator $L$

$$x(\alpha) = \arg\min_x \{\|b - A_k x\|^2 + \alpha^2 \|Lx\|^2\}$$

Solve with standard form if $L$ invertible.

Desirable automatic estimation of $\alpha$

How to efficiently solve and find $\alpha^{\text{opt}}$ for large scale problems?
Tikhonov Regularization: regularization parameter $\alpha$

Filter Functions

$$\gamma_i(\alpha) = \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2}, \quad i = 1 \ldots r,$$

Solves Standard Form

$$x(\alpha) = \arg\min_x \{ \|b - Ax\|^2 + \alpha^2 \|x\|^2 \}$$

$$\approx \arg\min_x \{ \|b - A_kx\|^2 + \alpha^2 \|x\|^2 \}$$

Generalized Tikhonov has operator $L$

$$x(\alpha) = \arg\min_x \{ \|b - A_kx\|^2 + \alpha^2 \|Lx\|^2 \}$$

Solve with standard form if $L$ invertible.

Desirable automatic estimation of $\alpha$

How to efficiently solve and find $\alpha^{opt}$ for large scale problems?
Tikhonov Regularization: regularization parameter $\alpha$

**Filter Functions**

$$\gamma_i(\alpha) = \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2}, \ i = 1 \ldots r,$$

**Solves Standard Form**

$$x(\alpha) = \arg\min_x \{\|b - Ax\|^2 + \alpha^2\|x\|^2\}$$

$$\approx \arg\min_x \{\|b - A_kx\|^2 + \alpha^2\|x\|^2\}$$

**Generalized Tikhonov has operator $L$**

$$x(\alpha) = \arg\min_x \{\|b - A_kx\|^2 + \alpha^2\|Lx\|^2\}$$

Solve with standard form if $L$ invertible.

**Desirable automatic estimation of $\alpha$**

How to efficiently solve and find $\alpha^{\text{opt}}$ for large scale problems?
Tikhonov Regularization: regularization parameter $\alpha$

Filter Functions

$$\gamma_i(\alpha) = \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2}, \quad i = 1 \ldots r,$$

Solves Standard Form

$$\mathbf{x}(\alpha) = \underset{\mathbf{x}}{\arg\min} \left\{ \| \mathbf{b} - A\mathbf{x} \|^2 + \alpha^2 \| \mathbf{x} \|^2 \right\}$$

$$\approx \underset{\mathbf{x}}{\arg\min} \left\{ \| \mathbf{b} - A_k\mathbf{x} \|^2 + \alpha^2 \| \mathbf{x} \|^2 \right\}$$

Generalized Tikhonov has operator $L$

$$\mathbf{x}(\alpha) = \underset{\mathbf{x}}{\arg\min} \left\{ \| \mathbf{b} - A_k\mathbf{x} \|^2 + \alpha^2 \| L\mathbf{x} \|^2 \right\}$$

Solve with standard form if $L$ invertible.

Desirable automatic estimation of $\alpha$

How to efficiently solve and find $\alpha^{\text{opt}}$ for large scale problems?
Tikhonov Regularization: regularization parameter $\alpha$

Filter Functions

$$\gamma_i(\alpha) = \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2}, \ i = 1 \ldots r,$$

Solves Standard Form

$$x(\alpha) = \arg\min_x \left\{ \|b - Ax\|^2 + \alpha^2 \|x\|^2 \right\}$$

$$\approx \arg\min_x \left\{ \|b - A_kx\|^2 + \alpha^2 \|x\|^2 \right\}$$

Generalized Tikhonov has operator $L$

$$x(\alpha) = \arg\min_x \left\{ \|b - A_kx\|^2 + \alpha^2 \|Lx\|^2 \right\}$$

Solve with standard form if $L$ invertible.

Desirable automatic estimation of $\alpha$

How to efficiently solve and find $\alpha^{\text{opt}}$ for large scale problems?
Regularization Parameter Estimation: Minimize $F(\alpha)$ some $F$

Introduce $\phi_i(\alpha) = \frac{\alpha^2}{\alpha^2 + \sigma_i^2} = 1 - \gamma_i(\alpha)$, $i = 1 : r$, $\gamma_i = 0$, $i > k$.

Unbiased Predictive Risk: Minimize functional noise level $\eta^2$

$$U_k(\alpha) = \sum_{i=1}^{k} \phi_i^2(\alpha)(u_i^T b)^2 - 2\eta^2 \sum_{i=1}^{k} \phi_i(\alpha)$$

GCV: Minimize rational function $m^* = \min\{m, n\}$

$$G(\alpha) = \frac{\left(\sum_{i=1}^{m^*} \phi_i^2(\alpha)(u_i^T b)^2\right)}{\left(\sum_{i=1}^{m^*} \phi_i^2(\alpha)\right)}$$

WGCV: Minimize

$$WG(\alpha) = \frac{\left(\sum_{i=1}^{m^*} \phi_i^2(\alpha)(u_i^T b)^2\right)}{\left(1 + k - \omega k + \omega \sum_{i=1}^{k} \phi_i^2(\alpha)\right)}$$

How does $\alpha^{\text{opt}} = \arg\min F(\alpha)$ depend on $k$?
Regularization Parameter Estimation: Minimize $F(\alpha)$ some $F$

Introduce $\phi_i(\alpha) = \frac{\alpha^2}{\alpha^2 + \sigma_i^2} = 1 - \gamma_i(\alpha)$, $i = 1 : r$, $\gamma_i = 0$, $i > k$.

**Unbiased Predictive Risk** : Minimize functional noise level $\eta^2$

$$U_k(\alpha) = \sum_{i=1}^{k} \phi_i^2(\alpha)(u_i^T b)^2 - 2\eta^2 \sum_{i=1}^{k} \phi_i(\alpha)$$

**GCV** : Minimize rational function $m^* = \min\{m, n\}$

$$G(\alpha) = \frac{\left(\sum_{i=1}^{m^*} \phi_i^2(\alpha)(u_i^T b)^2\right)}{\left(\sum_{i=1}^{m^*} \phi_i^2(\alpha)\right)}$$

**WGCV** : Minimize

$$WG(\alpha) = \frac{\left(\sum_{i=1}^{m^*} \phi_i^2(\alpha)(u_i^T b)^2\right)}{\left(1 + k - \omega k + \omega \sum_{i=1}^{k} \phi_i^2(\alpha)\right)}$$

How does $\alpha_{opt} = \arg\min F(\alpha)$ depend on $k$?
Regularization Parameter Estimation: Minimize $F(\alpha)$ some $F$

Introduce $\phi_i(\alpha) = \frac{\alpha^2}{\alpha^2 + \sigma_i^2} = 1 - \gamma_i(\alpha)$, $i = 1 : r$, $\gamma_i = 0$, $i > k$.

**Unbiased Predictive Risk**: Minimize functional noise level $\eta^2$

$$U_k(\alpha) = \sum_{i=1}^{k} \phi_i^2(\alpha)(u_i^Tb)^2 - 2\eta^2 \sum_{i=1}^{k} \phi_i(\alpha)$$

**GCV**: Minimize rational function $m^* = \min\{m, n\}$

$$G(\alpha) = \frac{\left(\sum_{i=1}^{m^*} \phi_i^2(\alpha)(u_i^Tb)^2\right)}{\left(\sum_{i=1}^{m^*} \phi_i^2(\alpha)\right)}$$

**WGCV**: Minimize

$$WG(\alpha) = \frac{\left(\sum_{i=1}^{m^*} \phi_i^2(\alpha)(u_i^Tb)^2\right)}{\left(1 + k - \omega k + \omega \sum_{i=1}^{k} \phi_i^2(\alpha)\right)}$$

**How does $\alpha^{opt} = \arg\min F(\alpha)$ depend on $k$?**
Regularization Parameter Estimation: Minimize $F(\alpha)$ some $F$

Introduce $\phi_i(\alpha) = \frac{\alpha^2}{\alpha^2 + \sigma_i^2} = 1 - \gamma_i(\alpha)$, $i = 1 : r$, $\gamma_i = 0$, $i > k$.

Unbiased Predictive Risk: Minimize functional noise level $\eta^2$

$$U_k(\alpha) = \sum_{i=1}^{k} \phi_i^2(\alpha) (u_i^T b)^2 - 2\eta^2 \sum_{i=1}^{k} \phi_i(\alpha)$$

GCV: Minimize rational function $m^* = \min \{ m, n \}$

$$G(\alpha) = \frac{\left( \sum_{i=1}^{m^*} \phi_i^2(\alpha)(u_i^T b)^2 \right)}{\left( \sum_{i=1}^{m^*} \phi_i^2(\alpha) \right)}$$

WGCV: Minimize

$$WG(\alpha) = \frac{\left( \sum_{i=1}^{m^*} \phi_i^2(\alpha)(u_i^T b)^2 \right)}{\left( 1 + k - \omega k + \omega \sum_{i=1}^{k} \phi_i^2(\alpha) \right)}$$

How does $\alpha^{opt} = \arg\min F(\alpha)$ depend on $k$?
Regularization Parameter Estimation: Minimize $F(\alpha)$ some $F$

Introduce $\phi_i(\alpha) = \frac{\alpha^2}{\alpha^2 + \sigma_i^2} = 1 - \gamma_i(\alpha)$, $i = 1 : r$, $\gamma_i = 0$, $i > k$.

**Unbiased Predictive Risk**: Minimize functional noise level $\eta^2$

$$U_k(\alpha) = \sum_{i=1}^{k} \phi_i^2(\alpha) (u_i^T b)^2 - 2\eta^2 \sum_{i=1}^{k} \phi_i(\alpha)$$

**GCV**: Minimize rational function $m^* = \min\{m, n\}$

$$G(\alpha) = \frac{\left(\sum_{i=1}^{m^*} \phi_i^2(\alpha) (u_i^T b)^2\right)}{\left(\sum_{i=1}^{m^*} \phi_i^2(\alpha)\right)}$$

**WGCV**: Minimize

$$WG(\alpha) = \frac{\left(\sum_{i=1}^{m^*} \phi_i^2(\alpha) (u_i^T b)^2\right)}{\left(1 + k - \omega k + \omega \sum_{i=1}^{k} \phi_i^2(\alpha)\right)}$$

**How does $\alpha^{opt} = \arg\min F(\alpha)$ depend on $k$?**
Examples: $F(\alpha)$ Increasing truncation $k$. Noise level $\eta^2 = .0001$

Regularization parameter independent of $k$: unique minimum: $U_k(\alpha)$ increases.
Examples: $F(\alpha)$ Increasing truncation $k$. Noise level $\eta^2 = .0001$

Regularization parameter independent of $k$: unique minimum: 
$U_k(\alpha)$ increases.
Examples: $F(\alpha)$ Increasing truncation $k$. Noise level $\eta^2 = .0001$

GCV function $G_k(\alpha)$ with increasing $k$:

Regularization parameter dependent on $k$: nonunique minimum $W_k(\alpha)$ generally decreases
Examples: $F(\alpha)$ Increasing truncation $k$. Noise level $\eta^2 = .0001$

Regularization parameter dependent on $k$: nonunique minimum $W_k(\alpha)$ generally decreases
Examples: $F(\alpha)$ Increasing truncation $k$. Noise level $\eta^2 = .0001$

Regularization parameter dependent on $k$: unique minimum. $WG_k(\alpha)$ decreases
Examples: $F(\alpha)$ Increasing truncation $k$. Noise level $\eta^2 = .0001$

Regularization parameter dependent on $k$: unique minimum. $WG_k(\alpha)$ decreases
Coefficients of the data : \( \hat{b}_i = u_i^T b \)

Assumption : There exists \( \ell \) such that \( \mathbb{E}(\hat{b}_i^2) = \eta^2 \) for all \( i > \ell \), i.e. coefficients \( \hat{b}_i \) noise contaminated \( i > \ell \).

Define

\[
\alpha^\text{opt} = \operatorname{argmin} U(\alpha) \quad \text{and} \quad \alpha_k = \operatorname{argmin} U_k(\alpha),
\]

Theorem (Convergence of \( \alpha_k \) for UPRE)

Suppose that \( k = \ell + p, p > 0 \), then the sequence \( \{\alpha_k\} \) is on the average increasing with \( \lim_{k \to r} \alpha_k = \alpha^\text{opt} \). Furthermore \( \{U_k(\alpha_k)\} \) is increasing, with \( \lim_{k \to r} U_k(\alpha_k) = U(\alpha^\text{opt}) \).
Theorem on UPRE for the FTSVD regularization

Coefficients of the data: \( \hat{b}_i = u_i^T b \)

Assumption: There exists \( \ell \) such that \( \mathbb{E}(\hat{b}_i^2) = \eta^2 \) for all \( i > \ell \), i.e. coefficients \( \hat{b}_i \) noise contaminated \( i > \ell \).

Define

\[ \alpha^{\text{opt}} = \arg\min U(\alpha) \quad \text{and} \quad \alpha_k = \arg\min U_k(\alpha), \]

Theorem (Convergence of \( \alpha_k \) for UPRE)

Suppose that \( k = \ell + p, \ p > 0 \), then the sequence \( \{\alpha_k\} \) is on the average increasing with \( \lim_{k \to r} \alpha_k = \alpha^{\text{opt}} \). Furthermore \( \{U_k(\alpha_k)\} \) is increasing, with \( \lim_{k \to r} U_k(\alpha_k) = U(\alpha^{\text{opt}}) \).
Theorem on UPRE for the FTSVD regularization

Coefficients of the data : \( \hat{b}_i = u_i^T b \)

Assumption : There exists \( \ell \) such that \( E(\hat{b}_i^2) = \eta^2 \) for all \( i > \ell \), i.e. coefficients \( \hat{b}_i \) noise contaminated \( i > \ell \).

Define

\[
\alpha^{\text{opt}} = \arg\min U(\alpha) \quad \text{and} \quad \alpha_k = \arg\min U_k(\alpha),
\]

Theorem (Convergence of \( \alpha_k \) for UPRE)

Suppose that \( k = \ell + p, \ p > 0 \), then the sequence \( \{\alpha_k\} \) is on the average increasing with \( \lim_{k \to r} \alpha_k = \alpha^{\text{opt}} \). Furthermore \( \{U_k(\alpha_k)\} \) is increasing, with \( \lim_{k \to r} U_k(\alpha_k) = U(\alpha^{\text{opt}}) \).
Theorem on UPRE for the FTSVD regularization

Coefficients of the data: \( \hat{b}_i = u_i^T b \)

Assumption: There exists \( \ell \) such that \( \mathbb{E}(\hat{b}_i^2) = \eta^2 \) for all \( i > \ell \), i.e. coefficients \( \hat{b}_i \) noise contaminated \( i > \ell \).

Define

\[
\alpha^{\text{opt}} = \arg\min U(\alpha) \quad \text{and} \quad \alpha_k = \arg\min U_k(\alpha),
\]

Theorem (Convergence of \( \alpha_k \) for UPRE)

Suppose that \( k = \ell + p, p > 0 \), then the sequence \( \{\alpha_k\} \) is on the average increasing with \( \lim_{k \to r} \alpha_k = \alpha^{\text{opt}} \). Furthermore \( \{U_k(\alpha_k)\} \) is increasing, with \( \lim_{k \to r} U_k(\alpha_k) = U(\alpha^{\text{opt}}) \).
Theorem on WGCV for the FTSVD regularization

Weight parameter : $\omega = \frac{(k + 1)}{m^*}$

Define

$$\alpha^{\text{opt}} = \arg\min WG(\alpha) \quad \text{and} \quad \alpha_k = \arg\min WG_k(\alpha),$$

Theorem (Convergence of $\alpha_k$ for WGCV)

Suppose that $\omega = \frac{(k + 1)}{m^*}$, then $\{WG_k(\alpha_k)\}$ is decreasing, with $\lim_{k \to r} WG_k(\alpha_k) = WG(\alpha^{\text{opt}})$. 
Weight parameter: \( \omega = (k + 1)/m^* \)

Define

\[ \alpha^{\text{opt}} = \arg\min WG(\alpha) \quad \text{and} \quad \alpha_k = \arg\min WG_k(\alpha), \]

Theorem (Convergence of \( \alpha_k \) for WGCV)

Suppose that \( \omega = (k + 1)/m^* \), then \( \{WG_k(\alpha_k)\} \) is decreasing, with \( \lim_{k \to r} WG_k(\alpha_k) = WG(\alpha^{\text{opt}}) \).
Theorem on WGCV for the FTSVD regularization

Weight parameter : \( \omega = \frac{(k + 1)}{m^*} \)

Define

\[ \alpha_{opt} = \arg\min WG(\alpha) \quad \text{and} \quad \alpha_k = \arg\min WG_k(\alpha), \]

Theorem (Convergence of \( \alpha_k \) for WGCV)

Suppose that \( \omega = \frac{(k + 1)}{m^*} \), then \( \{WG_k(\alpha_k)\} \) is decreasing, with \( \lim_{k \to r} WG_k(\alpha_k) = WG(\alpha_{opt}) \).
Truncated Singular Value Decomposition approximates $A$

Observations:
1. Find $\alpha_k$ for TSVD with $k$ terms.
2. Determine optimal $k$ as $\alpha_k$ converges to $\alpha^{\text{opt}}$

Method approximating SVD with regularization: UPRE, WGCV

Regularizes the full problem

But finding the TSVD for large problems is not feasible
Truncated Singular Value Decomposition approximates $A$

Observations:

1. Find $\alpha_k$ for TSVD with $k$ terms.
2. Determine optimal $k$ as $\alpha_k$ converges to $\alpha^{opt}$

Method approximating SVD with regularization: UPRE, WGCV

Regularizes the full problem

But finding the TSVD for large problems is not feasible
Truncated Singular Value Decomposition approximates $A$

Observations:

1. Find $\alpha_k$ for TSVD with $k$ terms.
2. Determine optimal $k$ as $\alpha_k$ converges to $\alpha^{opt}$

Method approximating SVD with regularization: UPRE, WGCV

Regularizes the full problem

But finding the TSVD for large problems is not feasible
Truncated Singular Value Decomposition approximates $A$

Observations:
1. Find $\alpha_k$ for TSVD with $k$ terms.
2. Determine optimal $k$ as $\alpha_k$ converges to $\alpha^{\text{opt}}$.

Method approximating SVD with regularization: UPRE, WGCV

Regularizes the full problem

But finding the TSVD for large problems is not feasible.
Truncated Singular Value Decomposition approximates $A$

Observations:

1. Find $\alpha_k$ for TSVD with $k$ terms.
2. Determine optimal $k$ as $\alpha_k$ converges to $\alpha^{opt}$

Method approximating SVD with regularization: UPRE, WGCV

Regularizes the full problem

But finding the TSVD for large problems is not feasible
Large Scale - The LSQR iteration: Given $k$ defines range space

**LSQR** Let $\beta_1 := \|b\|_2$, and $e_1^{(k+1)}$ first column of $I_{k+1}$

Generate, lower bidiagonal $B_k \in \mathcal{R}^{(k+1) \times k}$, column orthonormal $H_{k+1} \in \mathcal{R}^{m \times (k+1)}$, $G_k \in \mathcal{R}^{n \times k}$

$$AG_k = H_{k+1}B_k, \quad \beta_1 H_{k+1}e_1^{(k+1)} = b.$$ 

Projected Problem on projected space: (standard Tikhonov)

$$w_k(\zeta_k) = \arg\min_{w \in \mathcal{R}^k} \left\{ \|B_k w - \beta_1 e_1^{(k+1)}\|_2^2 + \zeta_k^2 \|w\|_2^2 \right\}.$$ 

Projected Solution depends on $\zeta_k^{opt}$: Let $B_k = \tilde{U}\tilde{\Sigma}\tilde{V}^T$

$$x_k(\zeta_k^{opt}) = G_k w_k(\zeta_k^{opt}) = \beta_1 G_k \sum_{i=1}^{k+1} \gamma_i(\zeta_k^{opt}) \frac{\tilde{u}_i^T e_1^{(k+1)}}{\tilde{\sigma}_i} \tilde{v}_i$$

$$= \sum_{i=1}^{k} \gamma_i(\zeta_k^{opt}) \frac{\tilde{u}_i^T (H_{k+1}b)}{\tilde{\sigma}_i} G_k \tilde{v}_i$$

Approximate SVD: $\tilde{A}_k = (H_{k+1}\tilde{U})\tilde{\Sigma}(G_k\tilde{V})^T$
**Large Scale - The LSQR iteration: Given** \( k \) defines range space

**LSQR** Let \( \beta_1 := \|b\|_2 \), and \( e_1^{(k+1)} \) first column of \( I_{k+1} \)

Generate, lower bidiagonal \( B_k \in \mathcal{R}^{(k+1) \times k} \), column orthonormal \( H_{k+1} \in \mathcal{R}^{m \times (k+1)} \), \( G_k \in \mathcal{R}^{n \times k} \)

\[
AG_k = H_{k+1} B_k, \quad \beta_1 H_{k+1} e_1^{(k+1)} = b.
\]

**Projected Problem** on projected space: (standard Tikhonov)

\[
w_k(\zeta_k) = \arg\min_{w \in \mathcal{R}^k} \{\|B_k w - \beta_1 e_1^{(k+1)}\|_2^2 + \zeta_k^2 \|w\|_2^2\}.
\]

**Projected Solution** depends on \( \zeta_{k}^{\text{opt}} \): Let \( B_k = \tilde{U} \tilde{\Sigma} \tilde{V}^T \)

\[
x_k(\zeta_{k}^{\text{opt}}) = G_k w_k(\zeta_{k}^{\text{opt}}) = \beta_1 G_k \sum_{i=1}^{k+1} \gamma_i(\zeta_{k}^{\text{opt}}) \frac{\tilde{u}_i^T e_1^{(k+1)}}{\tilde{\sigma}_i} \tilde{v}_i
\]

\[
= \sum_{i=1}^{k} \gamma_i(\zeta_{k}^{\text{opt}}) \frac{\tilde{u}_i^T (H_{k+1}^T b)}{\tilde{\sigma}_i} G_k \tilde{v}_i
\]

**Approximate SVD:** \( \tilde{A}_k = (H_{k+1} \tilde{U}) \tilde{\Sigma} (G_k \tilde{V})^T \)
Large Scale - The LSQR iteration: Given \( k \) defines range space

**LSQR**  Let \( \beta_1 := \|b\|_2 \), and \( e_1^{(k+1)} \) first column of \( I_{k+1} \)
Generate, lower bidiagonal \( B_k \in \mathcal{R}^{(k+1) \times k} \), column orthonormal \( H_{k+1} \in \mathcal{R}^{m \times (k+1)} \), \( G_k \in \mathcal{R}^{n \times k} \)

\[
AG_k = H_{k+1}B_k, \quad \beta_1 H_{k+1}e_1^{(k+1)} = b.
\]

**Projected Problem**  on projected space: (standard Tikhonov)

\[
w_k(\zeta_k) = \arg\min_{w \in \mathcal{R}^k} \{\|B_k w - \beta_1 e_1^{(k+1)}\|_2^2 + \zeta_k^2 \|w\|_2^2\}.
\]

**Projected Solution**  depends on \( \zeta_{\text{opt}}^k \): Let \( B_k = \tilde{U}\tilde{\Sigma}\tilde{V}^T \)

\[
x_k(\zeta_{\text{opt}}^k) = G_k w_k(\zeta_{\text{opt}}^k) = \beta_1 G_k \sum_{i=1}^{k+1} \gamma_i(\zeta_{\text{opt}}^k) \frac{\tilde{u}_i^T e_1^{(k+1)}}{\tilde{\sigma}_i} \tilde{v}_i
\]

\[
= \sum_{i=1}^{k} \gamma_i(\zeta_{\text{opt}}^k) \frac{\tilde{u}_i^T (H_{k+1}^T b)}{\tilde{\sigma}_i} G_k \tilde{v}_i
\]

**Approximate SVD:** \( \tilde{A}_k = (H_{k+1}\tilde{U})\tilde{\Sigma}(G_k\tilde{V})^T \)
The LSQR iteration: Given \( k \) defines range space

**LSQR**  Let \( \beta_1 := \| b \|_2 \), and \( e_1^{(k+1)} \) first column of \( I_{k+1} \)

Generate, lower bidiagonal \( B_k \in \mathcal{R}^{(k+1)\times k} \), column orthonormal \( H_{k+1} \in \mathcal{R}^{m\times(k+1)} \), \( G_k \in \mathcal{R}^{n\times k} \)

\[
AG_k = H_{k+1}B_k, \quad \beta_1 H_{k+1} e_1^{(k+1)} = b.
\]

**Projected Problem**  on projected space: (standard Tikhonov)

\[
w_k(\zeta_k) = \arg\min_{w \in \mathcal{R}^k} \{ \| B_k w - \beta_1 e_1^{(k+1)} \|_2^2 + \zeta_k^2 \| w \|_2^2 \}.
\]

**Projected Solution**  depends on \( \zeta_{k}^{\text{opt}} \): Let \( B_k = \tilde{U} \tilde{\Sigma} \tilde{V}^T \)

\[
x_k(\zeta_{k}^{\text{opt}}) = G_k w_k(\zeta_{k}^{\text{opt}}) = \beta_1 G_k \sum_{i=1}^{k+1} \gamma_i(\zeta_{k}^{\text{opt}}) \tilde{u}_i^T \frac{e_1^{(k+1)}}{\tilde{\sigma}_i} \tilde{v}_i
\]

\[
= \sum_{i=1}^{k} \gamma_i(\zeta_{k}^{\text{opt}}) \frac{\tilde{u}_i^T (H_{k+1}^T b)}{\tilde{\sigma}_i} G_k \tilde{v}_i
\]

**Approximate SVD:** \( \tilde{A}_k = (H_{k+1} \tilde{U}) \tilde{\Sigma} (G_k \tilde{V})^T \)
Large Scale - The LSQR iteration: Given $k$ defines range space

**LSQR** Let $\beta_1 := \|b\|_2$, and $e_1^{(k+1)}$ first column of $I_{k+1}$

Generate, lower bidiagonal $B_k \in \mathcal{R}^{(k+1) \times k}$, column orthonormal $H_{k+1} \in \mathcal{R}^{m \times (k+1)}$, $G_k \in \mathcal{R}^{n \times k}$

$$AG_k = H_{k+1}B_k, \quad \beta_1H_{k+1}e_1^{(k+1)} = b.$$ 

**Projected Problem** on projected space: (standard Tikhonov)

$$w_k(\zeta_k) = \arg\min_{w \in \mathcal{R}^k} \{\|B_k w - \beta_1 e_1^{(k+1)}\|_2^2 + \zeta_k^2 \|w\|_2^2\}.$$ 

**Projected Solution** depends on $\zeta_k^{opt}$: Let $B_k = \tilde{U} \tilde{\Sigma} \tilde{V}^T$

$$x_k(\zeta_k^{opt}) = G_k w_k(\zeta_k^{opt}) = \beta_1 G_k \sum_{i=1}^{k+1} \gamma_i(\zeta_k^{opt}) \frac{\tilde{u}_i}{\tilde{\sigma}_i} e_1^{(k+1)} \tilde{v}_i$$

$$= \sum_{i=1}^{k} \gamma_i(\zeta_k^{opt}) \frac{\tilde{u}_i^T (H_{k+1} \tilde{b})}{\tilde{\sigma}_i} G_k \tilde{v}_i$$

**Approximate SVD:** $\tilde{A}_k = (H_{k+1} \tilde{U}) \tilde{\Sigma} (G_k \tilde{V})^T$
Large Scale - The LSQR iteration: Given $k$ defines range space

**LSQR** Let $\beta_1 := \|b\|_2$, and $e_1^{(k+1)}$ first column of $I_{k+1}$
Generate, lower bidiagonal $B_k \in \mathcal{R}^{(k+1) \times k}$, column orthonormal $H_{k+1} \in \mathcal{R}^{m \times (k+1)}$, $G_k \in \mathcal{R}^{n \times k}$

$$AG_k = H_{k+1} B_k, \quad \beta_1 H_{k+1} e_1^{(k+1)} = b.$$  

**Projected Problem** on projected space: (standard Tikhonov)

$$w_k(\zeta_k) = \arg\min_{w \in \mathcal{R}^k} \{ \|B_k w - \beta_1 e_1^{(k+1)}\|_2^2 + \zeta_k^2 \|w\|_2^2 \}.$$  

**Projected Solution** depends on $\zeta_k^{opt}$: Let $B_k = \tilde{U} \tilde{\Sigma} \tilde{V}^T$

$$x_k(\zeta_k^{opt}) = G_k w_k(\zeta_k^{opt}) = \beta_1 G_k \sum_{i=1}^{k+1} \gamma_i(\zeta_k^{opt}) \frac{\tilde{u}_i^T e_1^{(k+1)}}{\tilde{\sigma}_i} \tilde{v}_i$$

$$= \sum_{i=1}^k \gamma_i(\zeta_k^{opt}) \frac{\tilde{u}_i^T (H_{k+1}^T b)}{\tilde{\sigma}_i} G_k \tilde{v}_i$$

**Approximate SVD:** $\tilde{A}_k = (H_{k+1} \tilde{U}) \tilde{\Sigma} (G_k \tilde{V})^T$
Large Scale - The LSQR iteration: Given \( k \) defines range space

**LSQR** Let \( \beta_1 := \| b \|_2 \), and \( e_1^{(k+1)} \) first column of \( I_{k+1} \)

Generate, lower bidiagonal \( B_k \in \mathcal{R}^{(k+1) \times k} \), column orthonormal \( H_{k+1} \in \mathcal{R}^{m \times (k+1)} \), \( G_k \in \mathcal{R}^{n \times k} \)

\[
AG_k = H_{k+1}B_k, \quad \beta_1 H_{k+1}e_1^{(k+1)} = b.
\]

**Projected Problem** on projected space: (standard Tikhonov)

\[
w_k(\zeta_k) = \arg\min_{w \in \mathcal{R}^k} \{ \| B_k w - \beta_1 e_1^{(k+1)} \|_2^2 + \zeta_k^2 \| w \|_2^2 \}.
\]

**Projected Solution** depends on \( \zeta_k^{\text{opt}} \): Let \( B_k = \tilde{U}\tilde{\Sigma}\tilde{V}^T \)

\[
x_k(\zeta_k^{\text{opt}}) = G_k w_k(\zeta_k^{\text{opt}}) = \beta_1 G_k \sum_{i=1}^{k+1} \gamma_i(\zeta_k^{\text{opt}}) \frac{\tilde{u}_i^T e_1^{(k+1)}}{\tilde{\sigma}_i} \tilde{v}_i
\]

\[
= \sum_{i=1}^{k} \gamma_i(\zeta_k^{\text{opt}}) \frac{\tilde{u}_i^T (H_{k+1}^T b)}{\tilde{\sigma}_i} G_k \tilde{v}_i
\]

**Approximate SVD:** \( \tilde{A}_k = (H_{k+1}\tilde{U})\tilde{\Sigma}(G_k\tilde{V})^T \)
The Randomized Singular Value Decomposition: Proto [HMT11]

\[ A \in \mathcal{R}^{m \times n}, \text{ target rank } k, \text{ oversampling parameter } p, \]
\[ k + p = kp \ll m. \text{ Power factor } q. \text{ Compute } A \approx \overline{A}_k = \overline{U}_k \overline{\Sigma}_k \overline{V}_k^T, \]
\[ \overline{U}_k \in \mathcal{R}^{m \times k}, \overline{\Sigma}_k \in \mathcal{R}^{k \times k}, \overline{V}_k \in \mathcal{R}^{n \times k}. \]

1: Generate a Gaussian random matrix \( \Omega \in \mathcal{R}^{n \times kp}. \)
2: Compute \( Y = A\Omega \in \mathcal{R}^{m \times kp}. \) \( Y = \text{orth}(Y) \)
3: If \( q > 0 \) repeat \( q \) times \( \{Y = A(A^TY), Y = \text{orth}(Y)\} \). Power
4: Form \( B = Y^TA \in \mathcal{R}^{kp \times n}. \) \( (Q = Y) \)
5: Economy SVD \( B = U_B \Sigma_B V_B^T, U_B \in \mathcal{R}^{kp \times kp}, V_B \in \mathcal{R}^{k \times k} \)
6: \( \overline{U}_k = QU_B(:, 1 : k), \overline{V}_k = V_B(:, 1 : k), \overline{\Sigma}_k = \Sigma_B(1 : k, 1 : k) \)

Projected RSVD Problem

\[ x_k(\mu_k) = \arg\min_{x \in \mathcal{R}^k} \{ \| \overline{A}_k x - b \|_2^2 + \mu_k^2 \| x \|_2^2 \}. \]

\[ = \sum_{i=1}^{k} \gamma_i(\mu_k) (\overline{u}_k)_i^T \frac{b}{(\overline{\sigma}_k)_i} (\overline{v}_k)_i. \]

Approximate SVD \( \overline{A}_k = \overline{U}_k \overline{\Sigma}_k \overline{V}_k^T \)
The Randomized Singular Value Decomposition: Proto [HMT11]

\( A \in \mathbb{R}^{m \times n} \), target rank \( k \), oversampling parameter \( p \),

\( k + p = kp \ll m \). Power factor \( q \). Compute \( A \approx \overline{A}_k = U_k \Sigma_k V_k^T \),

\( U_k \in \mathbb{R}^{m \times k} \), \( \Sigma_k \in \mathbb{R}^{k \times k} \), \( V_k \in \mathbb{R}^{n \times k} \).

1: Generate a Gaussian random matrix \( \Omega \in \mathbb{R}^{n \times kp} \).
2: Compute \( Y = A \Omega \in \mathbb{R}^{m \times kp} \). \( Y = \text{orth}(Y) \)
3: If \( q > 0 \) repeat \( q \) times \( \{ Y = A(A^T Y), Y = \text{orth}(Y) \} \). Power
4: Form \( B = Y^T A \in \mathbb{R}^{kp \times n} \). \( (Q = Y) \)
5: Economy SVD \( B = U_B \Sigma_B V_B^T \), \( U_B \in \mathbb{R}^{kp \times kp} \), \( V_B \in \mathbb{R}^{k \times k} \)
6: \( U_k = QU_B(:,1:k) \), \( V_k = V_B(:,1:k) \), \( \Sigma_k = \Sigma_B(1:k,1:k) \)

Projected RSVD Problem

\[ x_k(\mu_k) = \arg\min_{x \in \mathbb{R}^k} \{ \| \overline{A}_k x - b \|_2^2 + \mu_k^2 \| x \|_2^2 \} . \]

\[ = \sum_{i=1}^k \gamma_i(\mu_k) \frac{(\overline{u}_k)_i}{(\overline{\sigma}_k)_i} (\overline{v}_k)_i . \]

Approximate SVD \( \overline{A}_k = U_k \Sigma_k V_k^T \)
The Randomized Singular Value Decomposition: Proto [HMT11]

\[ A \in \mathcal{R}^{m \times n} \], target rank \( k \), oversampling parameter \( p \),

\[ k + p = kp \ll m \]. Power factor \( q \). Compute \( A \approx A_k = \bar{U}_k \bar{\Sigma}_k \bar{V}_k^T \),

\[ \bar{U}_k \in \mathcal{R}^{m \times k}, \bar{\Sigma}_k \in \mathcal{R}^{k \times k}, \bar{V}_k \in \mathcal{R}^{n \times k}. \]

1: Generate a Gaussian random matrix \( \Omega \in \mathcal{R}^{n \times kp} \).

2: Compute \( Y = A \Omega \in \mathcal{R}^{m \times kp}. Y = \text{orth}(Y) \)

3: If \( q > 0 \) repeat \( q \) times \( \{ Y = A(A^T Y), Y = \text{orth}(Y) \} \). Power

4: Form \( B = Y^T A \in \mathcal{R}^{kp \times n}. (Q = Y) \)

5: Economy SVD \( B = U_B \Sigma_B V_B^T, U_B \in \mathcal{R}^{kp \times kp}, V_B \in \mathcal{R}^{k \times k} \)

6: \( \bar{U}_k = QU_B(:, 1 : k), \bar{V}_k = V_B(:, 1 : k), \bar{\Sigma}_k = \Sigma_B(1 : k, 1 : k) \)

Projected RSVD Problem

\[ x_k(\mu_k) = \arg\min_{x \in \mathcal{R}^k} \{ \| A_k x - b \|_2^2 + \mu_k^2 \| x \|_2^2 \}. \]

\[ = \sum_{i=1}^{k} \gamma_i(\mu_k) \frac{(\bar{u}_k)_{i}^T b}{(\bar{\sigma}_k)_{i}} (\bar{v}_k)_{i}. \]

Approximate SVD \( \overline{A}_k = \bar{U}_k \bar{\Sigma}_k \bar{V}_k^T \).
\( A \in \mathcal{R}^{m \times n} \), target rank \( k \), oversampling parameter \( p \),
\[ k + p = kp \ll m. \] Power factor \( q \). Compute \( A \approx \bar{A}_k = \bar{U}_k \bar{\Sigma}_k \bar{V}_k^T \),
\( \bar{U}_k \in \mathcal{R}^{m \times k}, \bar{\Sigma}_k \in \mathcal{R}^{k \times k}, \bar{V}_k \in \mathcal{R}^{n \times k} \).

1: Generate a Gaussian random matrix \( \Omega \in \mathcal{R}^{n \times kp} \).
2: Compute \( Y = A\Omega \in \mathcal{R}^{m \times kp}. \ Y = \text{orth}(Y) \)
3: If \( q > 0 \) repeat \( q \) times \( \{Y = A(A^TY), Y = \text{orth}(Y)\} \). Power
4: Form \( B = Y^T A \in \mathcal{R}^{kp \times n}. \) (\( Q = Y \))
5: Economy SVD \( B = U_B \Sigma_B V_B^T, U_B \in \mathcal{R}^{kp \times kp}, V_B \in \mathcal{R}^{k \times k} \)
6: \( \bar{U}_k = QU_B(:, 1 : k), \bar{V}_k = V_B(:, 1 : k), \bar{\Sigma}_k = \Sigma_B(1 : k, 1 : k) \)

Projected RSVD Problem

\[ x_k(\mu_k) = \arg\min_{x \in \mathcal{R}^k} \{\|\bar{A}_k x - b\|_2^2 + \mu_k^2 \|x\|_2^2\}. \]

\[ = \sum_{i=1}^{k} \gamma_i(\mu_k) \frac{(\bar{u}_k)_i^T b}{(\bar{\sigma}_k)_i} (\bar{v}_k)_i. \]

Approximate SVD \( \bar{A}_k = \bar{U}_k \bar{\Sigma}_k \bar{V}_k^T \)
$A \in \mathcal{R}^{m \times n}$, target rank $k$, oversampling parameter $p$, $k + p = kp \ll m$. Power factor $q$. Compute $A \approx \overline{A}_k = \overline{U}_k \overline{\Sigma}_k \overline{V}_k^T$, 
$\overline{U}_k \in \mathcal{R}^{m \times k}$, $\overline{\Sigma}_k \in \mathcal{R}^{k \times k}$, $\overline{V}_k \in \mathcal{R}^{n \times k}$.

1: Generate a Gaussian random matrix $\Omega \in \mathcal{R}^{n \times kp}$.
2: Compute $Y = A\Omega \in \mathcal{R}^{m \times kp}$. $Y = \text{orth}(Y)$
3: If $q > 0$ repeat $q$ times \{ $Y = A(A^T Y)$, $Y = \text{orth}(Y)$ \}. Power
4: Form $B = Y^T A \in \mathcal{R}^{kp \times n}$. ($Q = Y$)
5: Economy SVD $B = U_B \Sigma_B V_B^T$, $U_B \in \mathcal{R}^{kp \times kp}$, $V_B \in \mathcal{R}^{k \times k}$
6: $\overline{U}_k = QU_B(:, 1 : k)$, $\overline{V}_k = V_B(:, 1 : k)$, $\overline{\Sigma}_k = \Sigma_B(1 : k, 1 : k)$

Projected RSVD Problem

$$x_k(\mu_k) = \arg\min_{x \in \mathcal{R}^k} \{ \| \overline{A}_k x - b \|^2_2 + \mu_k^2 \| x \|^2_2 \}. $$

$$= \sum_{i=1}^k \gamma_i(\mu_k) \frac{(\overline{u}_k)_i^T b}{(\overline{\sigma}_k)_i} (\overline{v}_k)_i. $$

Approximate SVD $\overline{A}_k = \overline{U}_k \overline{\Sigma}_k \overline{V}_k^T$
$A \in \mathcal{R}^{m \times n}$, target rank $k$, oversampling parameter $p$, $k + p = kp \ll m$. Power factor $q$. Compute $A \approx \overline{A}_k = U_k \Sigma_k V_k^T$, $U_k \in \mathcal{R}^{m \times k}$, $\Sigma_k \in \mathcal{R}^{k \times k}$, $V_k \in \mathcal{R}^{n \times k}$.

1: Generate a Gaussian random matrix $\Omega \in \mathcal{R}^{n \times kp}$.
2: Compute $Y = A\Omega \in \mathcal{R}^{m \times kp}$. $Y = \text{orth}(Y)$
3: If $q > 0$ repeat $q$ times $\{Y = A(A^TY), Y = \text{orth}(Y)\}$. Power
4: Form $B = Y^TA \in \mathcal{R}^{kp \times n}$. ($Q = Y$)
5: Economy SVD $B = U_B \Sigma_B V_B^T$, $U_B \in \mathcal{R}^{kp \times kp}$, $V_B \in \mathcal{R}^{k \times k}$
6: $U_k = QU_B(:, 1 : k)$, $V_k = V_B(:, 1 : k)$, $\Sigma_k = \Sigma_B(1 : k, 1 : k)$

Projected RSVD Problem

$$x_k(\mu_k) = \arg\min_{x \in \mathcal{R}^k} \{\|\overline{A}_k x - b\|_2^2 + \mu_k^2 \|x\|_2^2\}.$$

$$= \sum_{i=1}^{k} \gamma_i(\mu_k) \frac{(u_k)^T_i b}{(\sigma_k)_i} (\overline{v}_k)_i.$$

Approximate SVD $\overline{A}_k = U_k \Sigma_k V_k^T$.
The Randomized Singular Value Decomposition: Proto [HMT11]

$A \in \mathbb{R}^{m \times n}$, target rank $k$, oversampling parameter $p$,

$k + p = kp \ll m$. Power factor $q$. Compute $A \approx \overline{A}_k = \overline{U}_k \overline{\Sigma}_k \overline{V}_k^T$,

$\overline{U}_k \in \mathbb{R}^{m \times k}, \overline{\Sigma}_k \in \mathbb{R}^{k \times k}, \overline{V}_k \in \mathbb{R}^{n \times k}$.

1: Generate a Gaussian random matrix $\Omega \in \mathbb{R}^{n \times kp}$.
2: Compute $Y = A\Omega \in \mathbb{R}^{m \times kp}$. $Y = \text{orth}(Y)$
3: If $q > 0$ repeat $q$ times $\{Y = A(A^TY), Y = \text{orth}(Y)\}$. Power
4: Form $B = Y^TA \in \mathbb{R}^{kp \times n}$. ($Q = Y$)
5: Economy SVD $B = U_B\Sigma_B V_B^T$, $U_B \in \mathbb{R}^{kp \times kp}$, $V_B \in \mathbb{R}^{k \times k}$
6: $\overline{U}_k = QU_B(:,1:k)$, $\overline{V}_k = V_B(:,1:k)$, $\overline{\Sigma}_k = \Sigma_B(1:k,1:k)$

Projected RSVD Problem

$$\begin{align*}
x_k(\mu_k) &= \arg\min_{x \in \mathbb{R}^k} \{\|\overline{A}_k x - b\|_2^2 + \mu_k^2\|x\|_2^2\}.
= \sum_{i=1}^k \gamma_i(\mu_k) \frac{(\overline{u}_k)_i^T b}{(\overline{\sigma}_k)_i} (\overline{v}_k)_i.
\end{align*}$$

Approximate SVD $\overline{A}_k = \overline{U}_k \overline{\Sigma}_k \overline{V}_k^T$. 
The Randomized Singular Value Decomposition: Proto [HMT11]

\( A \in \mathbb{R}^{m \times n} \), target rank \( k \), oversampling parameter \( p \),
\( k + p = kp \ll m \). Power factor \( q \). Compute \( A \approx \overline{A}_k = U_k \Sigma_k \overline{V}_k^T \),
\( \overline{U}_k \in \mathbb{R}^{m \times k} \), \( \Sigma_k \in \mathbb{R}^{k \times k} \), \( \overline{V}_k \in \mathbb{R}^{n \times k} \).

1: Generate a Gaussian random matrix \( \Omega \in \mathbb{R}^{n \times kp} \).
2: Compute \( Y = A \Omega \in \mathbb{R}^{m \times kp} \). \( Y = \text{orth}(Y) \)
3: If \( q > 0 \) repeat \( q \) times \( \{ Y = A(A^T Y), Y = \text{orth}(Y) \} \). Power
4: Form \( B = Y^T A \in \mathbb{R}^{kp \times n} \). \( (Q = Y) \)
5: Economy SVD \( B = U_B \Sigma_B V_B^T \), \( U_B \in \mathbb{R}^{kp \times kp} \), \( V_B \in \mathbb{R}^{k \times k} \)
6: \( \overline{U}_k = QU_B(:, 1 : k) \), \( \overline{V}_k = V_B(:, 1 : k) \), \( \overline{\Sigma}_k = \Sigma_B(1 : k, 1 : k) \)

Projected RSVD Problem

\[
\begin{align*}
  x_k(\mu_k) &= \underset{x \in \mathbb{R}^k}{\text{argmin}} \{ \| \overline{A}_k x - b \|_2^2 + \mu_k^2 \| x \|_2^2 \} \\
  &= \sum_{i=1}^{k} \gamma_i(\mu_k) \frac{(\overline{u}_k)_i^T b}{(\overline{\sigma}_k)_i} (\overline{v}_k)_i.
\end{align*}
\]

Approximate SVD \( \overline{A}_k = \overline{U}_k \Sigma_k \overline{V}_k^T \)
$A \in \mathbb{R}^{m \times n}$, target rank $k$, oversampling parameter $p$,

$k + p = kp \ll m$. Power factor $q$. Compute $A \approx \overline{A}_k = \overline{U}_k \overline{\Sigma}_k \overline{V}_k^T$,

$\overline{U}_k \in \mathbb{R}^{m \times k}$, $\overline{\Sigma}_k \in \mathbb{R}^{k \times k}$, $\overline{V}_k \in \mathbb{R}^{n \times k}$.

1: Generate a Gaussian random matrix $\Omega \in \mathbb{R}^{n \times kp}$.
2: Compute $Y = A\Omega \in \mathbb{R}^{m \times kp}$. $Y = \text{orth}(Y)$
3: If $q > 0$ repeat $q$ times \{ $Y = A(A^TY)$, $Y = \text{orth}(Y)$ \}. Power
4: Form $B = Y^TA \in \mathbb{R}^{kp \times n}$. ($Q = Y$)
5: Economy SVD $B = U_B \Sigma_B V_B^T$, $U_B \in \mathbb{R}^{kp \times kp}, V_B \in \mathbb{R}^{k \times k}$
6: $\overline{U}_k = QU_B(:, 1 : k), \overline{V}_k = V_B(:, 1 : k), \overline{\Sigma}_k = \Sigma_B(1 : k, 1 : k)$

Projected RSVD Problem

\[
\mathbf{x}_k(\mu_k) = \arg\min_{\mathbf{x} \in \mathbb{R}^k} \left\{ \|\overline{A}_k \mathbf{x} - \mathbf{b}\|_2^2 + \mu_k^2 \|\mathbf{x}\|_2^2 \right\}.
\]

\[
= \sum_{i=1}^{k} \gamma_i(\mu_k) \frac{(\overline{u}_k)_i^T \mathbf{b}}{(\overline{\sigma}_k)_i} (\overline{v}_k)_i.
\]

Approximate SVD $\overline{A}_k = \overline{U}_k \overline{\Sigma}_k \overline{V}_k^T$. 
The Randomized Singular Value Decomposition: Proto [HMT11]

\[ A \in \mathcal{R}^{m \times n}, \text{ target rank } k, \text{ oversampling parameter } p, \]
\[ k + p = kp \ll m. \text{ Power factor } q. \text{ Compute } A \approx \overline{A}_k = \overline{U}_k \overline{\Sigma}_k \overline{V}_k^T, \]
\[ \overline{U}_k \in \mathcal{R}^{m \times k}, \overline{\Sigma}_k \in \mathcal{R}^{k \times k}, \overline{V}_k \in \mathcal{R}^{n \times k}. \]

1: Generate a Gaussian random matrix \( \Omega \in \mathcal{R}^{n \times kp} \).
2: Compute \( Y = A \Omega \in \mathcal{R}^{m \times kp} \). \( Y = \text{orth}(Y) \)
3: If \( q > 0 \) repeat \( q \) times \{ \( Y = A(A^T Y), Y = \text{orth}(Y) \} \). Power
4: Form \( B = Y^T A \in \mathcal{R}^{kp \times n} \). \( (Q = Y) \)
5: Economy SVD \( B = U_B \Sigma_B V_B^T, U_B \in \mathcal{R}^{kp \times kp}, V_B \in \mathcal{R}^{k \times k} \)
6: \( \overline{U}_k = QU_B(:,1:k), \overline{V}_k = V_B(:,1:k), \overline{\Sigma}_k = \Sigma_B(1:k,1:k) \)

Projected RSVD Problem

\[ x_k(\mu_k) = \underset{x \in \mathcal{R}^k}{\text{argmin}} \{ \| \overline{A}_k x - b \|^2_2 + \mu_k^2 \| x \|^2_2 \}. \]

\[ = \sum_{i=1}^k \gamma_i(\mu_k) \frac{(\overline{u}_k)_i^T b}{(\overline{\sigma}_k)_i} (\overline{v}_k)_i. \]

Approximate SVD \( \overline{A}_k = \overline{U}_k \overline{\Sigma}_k \overline{V}_k^T \)
### Summary Comparisons: rank \( k \)

<table>
<thead>
<tr>
<th></th>
<th>TSVD</th>
<th>LSQR</th>
<th>RSVD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>( A_k )</td>
<td>( \tilde{A}_k )</td>
<td>( \tilde{A}_k )</td>
</tr>
<tr>
<td>SVD</td>
<td>( U_k \Sigma_k V_k^T )</td>
<td>( (H_{k+1}{\tilde{U}})\Sigma(G_k{\tilde{V}})^T )</td>
<td>( U_k \Sigma_k V_k^T )</td>
</tr>
<tr>
<td>Terms</td>
<td>( \gamma_i(\alpha_k)\frac{u_i^T b}{\sigma_i} v_i )</td>
<td>( \gamma_i(\zeta_k)\frac{\tilde{u}<em>i^T (H</em>{k+1}^T b)}{\tilde{\sigma}_i} (G_k \tilde{v})_i )</td>
<td>( \gamma_i(\mu_k)\frac{(\tilde{u}_k)_i^T b}{(\tilde{\sigma}_k)_i} (\tilde{v}_k)_i )</td>
</tr>
<tr>
<td>Coeff</td>
<td>( \sigma_{k+1} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( | A - A_k | )</td>
<td></td>
<td>Theorem ( \tilde{A}_k )</td>
<td>Theorem ( \tilde{A}_k )</td>
</tr>
</tbody>
</table>

### Questions
1. Determine optimal \( k \) in each case?
2. Determine optimal \( \alpha_k, \zeta_k, \mu_k \). How are they related?
3. Do surrogate subproblems appropriately regularize the full problem?
4. Depends on approximation of singular space for \( A \)?
Summary Comparisons: rank $k$

<table>
<thead>
<tr>
<th></th>
<th>TSVD</th>
<th>LSQR</th>
<th>RSVD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVD</td>
<td>$A_k$</td>
<td>$\tilde{A}_k$</td>
<td>$\overline{A}_k$</td>
</tr>
<tr>
<td>Terms</td>
<td>$U_k \Sigma_k V_k^T$</td>
<td>$(H_{k+1} \tilde{U}) \tilde{\Sigma} (G_k \tilde{V})^T$</td>
<td>$U_k \Sigma_k V_k^T$</td>
</tr>
<tr>
<td>Coeff</td>
<td>$\gamma_i(\alpha_k) \frac{u_i^T b}{\sigma_i} v_i$</td>
<td>$\gamma_i(\zeta_k) \frac{\tilde{u}<em>i^T (H</em>{k+1}^T b)}{\tilde{\sigma}_i} (G_k \tilde{v})_i$</td>
<td>$\gamma_i(\mu_k) \frac{(\tilde{u}_k)^T b}{(\tilde{\sigma}_k)_i} (\tilde{v}_k)_i$</td>
</tr>
<tr>
<td>$|A - A_k|$</td>
<td>$\sigma_{k+1}$</td>
<td>Theorem $\tilde{A}_k$</td>
<td>Theorem $\overline{A}_k$</td>
</tr>
</tbody>
</table>

Questions

1. Determine optimal $k$ in each case?
2. Determine optimal $\alpha_k, \zeta_k, \mu_k$. How are they related?
3. Do surrogate subproblems appropriately regularize the full problem?
4. Depends on approximation of singular space for $A$?
### Summary Comparisons: rank $k$

<table>
<thead>
<tr>
<th></th>
<th>TSVD</th>
<th>LSQR</th>
<th>RSVD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>$A_k$</td>
<td>$	ilde{A}_k$</td>
<td>$	ilde{A}_k$</td>
</tr>
<tr>
<td>SVD</td>
<td>$U_k \Sigma_k V_k^T$</td>
<td>$(H_{k+1} \tilde{U}) \Sigma (G_k \tilde{V})^T$</td>
<td>$U_k \Sigma_k V_k^T$</td>
</tr>
<tr>
<td>Terms</td>
<td>$\gamma_i(\alpha_k) \frac{u_i^T b_i}{\sigma_i} v_i$</td>
<td>$\gamma_i(\zeta_k) \frac{\tilde{u}<em>i^T(H</em>{k+1}^T b_i)}{\tilde{\sigma}_i} (G_k \tilde{v})_i$</td>
<td>$\gamma_i(\mu_k) \frac{(\tilde{u}_k)_i^T b_i}{(\tilde{\sigma}_k)_i} (\tilde{v}_k)_i$</td>
</tr>
<tr>
<td>Coeff</td>
<td>$\sigma_{k+1}$</td>
<td>Theorem $\tilde{A}_k$</td>
<td>Theorem $\tilde{A}_k$</td>
</tr>
<tr>
<td>$</td>
<td></td>
<td>A - A_k</td>
<td></td>
</tr>
</tbody>
</table>

**Questions**

1. Determine optimal $k$ in each case?
2. Determine optimal $\alpha_k$, $\zeta_k$, $\mu_k$. How are they related?
3. Do surrogate subproblems appropriately regularize the full problem?
4. Depends on approximation of singular space for $A$?
### Summary Comparisons : rank $k$

<table>
<thead>
<tr>
<th></th>
<th>TSVD</th>
<th>LSQR</th>
<th>RSVD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>$A_k$</td>
<td>$\tilde{A}_k$</td>
<td>$\overline{A}_k$</td>
</tr>
<tr>
<td>SVD</td>
<td>$U_k \Sigma_k V_k^T$</td>
<td>$(H_{k+1} \tilde{U}) \tilde{\Sigma} (G_k \tilde{V})^T$</td>
<td>$U_k \Sigma_k \overline{V}_k^T$</td>
</tr>
<tr>
<td>Terms</td>
<td>$\gamma_i(\alpha_k) \frac{u_i^T b}{\sigma_i} v_i$</td>
<td>$\gamma_i(\zeta_k) \frac{\tilde{u}<em>i^T (H</em>{k+1} \tilde{b})}{\tilde{\sigma}_i} (G_k \tilde{v})_i$</td>
<td>$\gamma_i(\mu_k) \frac{(\overline{u}_k)^T b}{(\overline{\sigma})_i} (\overline{v}_k)_i$</td>
</tr>
<tr>
<td>Coeff</td>
<td>$\sigma_k + 1$</td>
<td>Theorem $\tilde{A}_k$</td>
<td>Theorem $\overline{A}_k$</td>
</tr>
<tr>
<td>$|A - A_k|$</td>
<td>$\sigma_k + 1$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Questions

1. Determine optimal $k$ in each case?
2. Determine optimal $\alpha_k$, $\zeta_k$, $\mu_k$. How are they related?
3. Do surrogate subproblems appropriately regularize the full problem?
4. Depends on approximation of singular space for $A$?
Summary Comparisons: rank $k$

<table>
<thead>
<tr>
<th>Model</th>
<th>TSVD</th>
<th>LSQR</th>
<th>RSVD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVD</td>
<td>$A_k$</td>
<td>$\tilde{A}_k$</td>
<td>$\overline{A}_k$</td>
</tr>
<tr>
<td>$U_k \Sigma_k V_k^T$</td>
<td>$(H_{k+1} \tilde{U}) \tilde{\Sigma}(G_k \tilde{V})^T$</td>
<td>$U_k \Sigma_k V_k^T$</td>
<td></td>
</tr>
<tr>
<td>Terms</td>
<td>$\gamma_i(\alpha_k) \frac{u_i^T b}{\sigma_i} v_i$</td>
<td>$\gamma_i(\tilde{\zeta}_k) \frac{\tilde{u}<em>i^T (H</em>{k+1}^T b)}{\tilde{\sigma}_i} (G_k \tilde{v})_i$</td>
<td>$\gamma_i(\mu_k) \frac{(\tilde{u}_k)_i^T b}{(\tilde{\sigma}_k)_i} (\tilde{v}_k)_i$</td>
</tr>
<tr>
<td>$|A - A_k|$</td>
<td>$\sigma_{k+1}$</td>
<td>Theorem $\tilde{A}_k$</td>
<td>Theorem $\overline{A}_k$</td>
</tr>
</tbody>
</table>

Questions

1. Determine optimal $k$ in each case?
2. Determine optimal $\alpha_k, \tilde{\zeta}_k, \mu_k$. How are they related?
3. Do surrogate subproblems appropriately regularize the full problem?
4. Depends on approximation of singular space for $A$?
### Summary Comparisons: rank $k$

<table>
<thead>
<tr>
<th></th>
<th>TSVD</th>
<th>LSQR</th>
<th>RSVD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>$A_k$</td>
<td>$\tilde{A}_k$</td>
<td>$\bar{A}_k$</td>
</tr>
<tr>
<td>SVD</td>
<td>$U_k \Sigma_k V_k^T$</td>
<td>$(H_{k+1} \tilde{U}) \tilde{\Sigma} (G_k \tilde{V})^T$</td>
<td>$U_k \bar{\Sigma}_k V_k^T$</td>
</tr>
<tr>
<td>Terms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coeff</td>
<td>$\gamma_i(\alpha_k) \frac{u_i^T b}{\sigma_i} v_i$</td>
<td>$\gamma_i(\zeta_k) \frac{\tilde{u}<em>i^T (H</em>{k+1}^T b)}{\tilde{\sigma}_i} (G_k \tilde{v})_i$</td>
<td>$\gamma_i(\mu_k) \frac{(\tilde{u}_k)_i^T b}{(\bar{\sigma}_k)_i} (\bar{v}_k)_i$</td>
</tr>
<tr>
<td>$|A - A_k|$</td>
<td>$\sigma_{k+1}$</td>
<td>Theorem $\tilde{A}_k$</td>
<td>Theorem $\bar{A}_k$</td>
</tr>
</tbody>
</table>

### Questions

1. Determine optimal $k$ in each case?
2. Determine optimal $\alpha_k, \zeta_k, \mu_k$. How are they related?
3. Do surrogate subproblems appropriately regularize the full problem?
4. Depends on approximation of singular space for $A$?
Define $\nu_k = \|A - \tilde{A}_k\| = \|\delta A_k\|$ then:

Theorem ([DDLT91]: (\(\sigma_i \neq \sigma_j\)) - looks at angles between singular vectors)

$u_i$ and $\tilde{u}_i$ ($v_i$ and $\tilde{v}_i$) are left (right) unit singular vectors of $A$ and $\tilde{A}_k$. For $\|\delta A_k\| \leq \nu_k$, if $2\nu_k < \min_{i \neq j} |\sigma_i - \sigma_j|$, then

$$\max(\sin \Theta(u_i, \tilde{u}_i), \sin \Theta(v_i, \tilde{v}_i)) \leq \frac{\nu_k}{\min_{i \neq j} |\sigma_i - \sigma_j| - \nu_k} \leq 1.$$ 

Theorem ([Jia16]: For fast decay of singular values near best rank $k$)

For $\ell$: $\hat{b}_\ell > \sigma_\ell$. Decay rate $\sigma_i = \zeta \rho^{-i}$, $\rho > 2$. Then

$\tilde{A}_k = H_{k+1}B_kG_k^T$ is a near best rank $k$ approximation to $A$ for $k = 1, 2, \ldots, \ell$. 
Theorems on $\tilde{A}_k$ Approximation of the spectral space: LSQR

Define $\nu_k = \|A - \tilde{A}_k\| = \|\delta A_k\|$ then:

Theorem ([DDLT91]: $(\sigma_i \neq \sigma_j)$ - looks at angles between singular vectors)

$u_i$ and $\tilde{u}_i$ ($v_i$ and $\tilde{v}_i$) are left (right) unit singular vectors of $A$ and $\tilde{A}_k$. For $\|\delta A_k\| \leq \nu_k$, if $2\nu_k < \min_{i \neq j} |\sigma_i - \sigma_j|$, then

$max(\sin \Theta(u_i, \tilde{u}_i), \sin \Theta(v_i, \tilde{v}_i)) \leq \frac{\nu_k}{\min_{i \neq j} |\sigma_i - \sigma_j| - \nu_k} \leq 1.$

Theorem ([Jia16]: For fast decay of singular values near best rank $k$)

For $\ell$: $\hat{b}_\ell > \sigma_\ell$. Decay rate $\sigma_i = \zeta \rho^{-i}$, $\rho > 2$. Then

$\tilde{A}_k = H_{k+1}B_k G_k^T$ is a near best rank $k$ approximation to $A$ for $k = 1, 2, \ldots, \ell$. 
Theorems on $\widetilde{A}_k$ Approximation of the spectral space: RSVD

**Theorem (Proto: near best rank $k$ approximation)**

*Target rank $k \geq 2$, oversampling $p \geq 2$, $k + p \leq \min\{m, n\}$.  

$$
E(\|A - QQ^T A\|) \leq \left[ 1 + \frac{4\sqrt{k + p}}{p - 1} \cdot \sqrt{\min\{m, n\}} \right] \sigma_{k+1}
$$

**Theorem (Power Iteration to force singular values to 0)**

$$
E(\|A - \overline{U}_k \Sigma_k \overline{V}_k^T\|) \leq \left( 1 + \left[ 1 + 4\sqrt{\frac{2\min\{m, n\}}{k - 1}} \right]^{1/(2q+1)} \right) \sigma_{k+1}
$$
Theorems on $\overline{A}_k$ Approximation of the spectral space: RSVD

**Theorem (Proto: near best rank $k$ approximation)**

*Target rank $k \geq 2$, oversampling $p \geq 2$, $k + p \leq \min\{m, n\}$.*

$$
E(\|A - QQ^T A\|) \leq \left[ 1 + \frac{4\sqrt{k + p}}{p - 1} \cdot \sqrt{\min\{m, n\}} \right] \sigma_{k+1}
$$

**Theorem (Power Iteration to force singular values to 0)**

$$
E(\|A - U_k \Sigma_k V_k^T \|) \leq \left( 1 + \left[ 1 + 4\sqrt{\frac{2\min\{m, n\}}{k - 1}} \right]^{1/(2q+1)} \right) \sigma_{k+1}
$$
Theorem ([RVA17]: with assumptions on the approximation of the spectral space for LSQR)

1. $\alpha^{\text{opt}}$ for $F^{\text{full}}(\alpha)$ can be estimated via LSQR
2. Minimizer of $F^{\text{proj}}(\zeta_k)$ is minimizer of $F^{\text{full}}(\zeta_k)$
3. $\zeta_k^{\text{opt}}$ depends on $k$, $\alpha^{\text{opt}}$ depends on $m^* =: \min(m, n)$
4. If $k^*$ approx numerical rank $A$, and right singular space is well-approximated $\zeta_{k^*}^{\text{opt}} \approx \alpha^{\text{opt}}$ for $K_{k^*}(A^T A, A^T b)$

Theorem (with assumptions on the approximation of the spectral space for RSVD follows from UPRE / WGCV convergence for FTSVD)

1. $\alpha^{\text{opt}}$ for $F^{\text{full}}(\alpha)$ can be estimated via RSVD
2. Minimizer of $F^{\text{rsvd}}(\mu_k)$ is minimizer of $F^{\text{full}}(\mu_k)$
3. $\mu_k^{\text{opt}}$ depends on $k$, $\alpha^{\text{opt}}$ depends on $m^* =: \min(m, n)$
4. For good rank $k$ approximation $\mu_k \approx \alpha_k$
Theorem ([RVA17]: with assumptions on the approximation of the spectral space for LSQR)

1. $\alpha^{\text{opt}}$ for $F^{\text{full}}(\alpha)$ can be estimated via LSQR
2. Minimizer of $F^{\text{proj}}(\zeta_k)$ is minimizer of $F^{\text{full}}(\zeta_k)$
3. $\zeta_k^{\text{opt}}$ depends on $k$, $\alpha^{\text{opt}}$ depends on $m^* =: \min(m, n)$
4. If $k^*$ approx numerical rank $A$, and right singular space is well-approximated $\zeta_{k^*}^{\text{opt}} \approx \alpha^{\text{opt}}$ for $K_k^*(A^TA, A^Tb)$

Theorem (with assumptions on the approximation of the spectral space for RSVD follows from UPRE / WGCV convergence for FTSVD)

1. $\alpha^{\text{opt}}$ for $F^{\text{full}}(\alpha)$ can be estimated via RSVD
2. Minimizer of $F^{\text{rsvd}}(\mu_k)$ is minimizer of $F^{\text{full}}(\mu_k)$
3. $\mu_k^{\text{opt}}$ depends on $k$, $\alpha^{\text{opt}}$ depends on $m^* =: \min(m, n)$
4. For good rank $k$ approximation $\mu_k \approx \alpha_k$
Regularization Estimation for UPRE / WGCV

Theorem ([RVA17]: with assumptions on the approximation of the spectral space for LSQR)

1. $\alpha^{\text{opt}}$ for $F^{\text{full}}(\alpha)$ can be estimated via LSQR
2. Minimizer of $F^{\text{proj}}(\zeta_k)$ is minimizer of $F^{\text{full}}(\zeta_k)$
3. $\zeta_k^{\text{opt}}$ depends on $k$, $\alpha^{\text{opt}}$ depends on $m^* =: \min(m, n)$
4. If $k^*$ approx numerical rank $A$, and right singular space is well-approximated $\zeta_{k^*}^{\text{opt}} \approx \alpha^{\text{opt}}$ for $K_{k^*}(A^T A, A^T b)$

Theorem (with assumptions on the approximation of the spectral space for RSVD follows from UPRE / WGCV convergence for FTSVD)

1. $\alpha^{\text{opt}}$ for $F^{\text{full}}(\alpha)$ can be estimated via RSVD
2. Minimizer of $F^{\text{rsvd}}(\mu_k)$ is minimizer of $F^{\text{full}}(\mu_k)$
3. $\mu_k^{\text{opt}}$ depends on $k$, $\alpha^{\text{opt}}$ depends on $m^* =: \min(m, n)$
4. For good rank $k$ approximation $\mu_k \approx \alpha_k$
Regularization Estimation for UPRE / WGCV

Theorem ([RVA17]: with assumptions on the approximation of the spectral space for LSQR)

1. $\alpha^\text{opt}$ for $F^{\text{full}}(\alpha)$ can be estimated via LSQR
2. Minimizer of $F^{\text{proj}}(\zeta_k)$ is minimizer of $F^{\text{full}}(\zeta_k)$
3. $\zeta_k^\text{opt}$ depends on $k$, $\alpha^\text{opt}$ depends on $m^* =: \min(m, n)$
4. If $k^*$ approx numerical rank $A$, and right singular space is well-approximated $\zeta_{k^*}^\text{opt} \approx \alpha^\text{opt}$ for $K_{k^*}(A^T A, A^T b)$

Theorem (with assumptions on the approximation of the spectral space for RSVD follows from UPRE / WGCV convergence for FTSVD)

1. $\alpha^\text{opt}$ for $F^{\text{full}}(\alpha)$ can be estimated via RSVD
2. Minimizer of $F^{\text{rsvd}}(\mu_k)$ is minimizer of $F^{\text{full}}(\mu_k)$
3. $\mu_k^\text{opt}$ depends on $k$, $\alpha^\text{opt}$ depends on $m^* =: \min(m, n)$
4. For good rank $k$ approximation $\mu_k \approx \alpha_k$
Regularization Estimation for UPRE / WGCV

Theorem ([RVA17]: with assumptions on the approximation of the spectral space for LSQR)

1. $\alpha^{\text{opt}}$ for $F^{\text{full}}(\alpha)$ can be estimated via LSQR
2. Minimizer of $F^{\text{proj}}(\zeta_k)$ is minimizer of $F^{\text{full}}(\zeta_k)$
3. $\zeta_k^{\text{opt}}$ depends on $k$, $\alpha^{\text{opt}}$ depends on $m^* =: \min(m, n)$
4. If $k^*$ approx numerical rank $A$, and right singular space is well-approximated $\zeta_{k^*}^{\text{opt}} \approx \alpha^{\text{opt}}$ for $K_{k^*}(A^T A, A^T b)$

Theorem (with assumptions on the approximation of the spectral space for RSVD follows from UPRE / WGCV convergence for FTSVD)

1. $\alpha^{\text{opt}}$ for $F^{\text{full}}(\alpha)$ can be estimated via RSVD
2. Minimizer of $F^{\text{rsvd}}(\mu_k)$ is minimizer of $F^{\text{full}}(\mu_k)$
3. $\mu_k^{\text{opt}}$ depends on $k$, $\alpha^{\text{opt}}$ depends on $m^* =: \min(m, n)$
4. For good rank $k$ approximation $\mu_k \approx \alpha_k$
Theorem ([RVA17]: with assumptions on the approximation of the spectral space for LSQR)

1. \( \alpha^{\text{opt}} \) for \( F^{\text{full}}(\alpha) \) can be estimated via LSQR
2. Minimizer of \( F^{\text{proj}}(\zeta_k) \) is minimizer of \( F^{\text{full}}(\zeta_k) \)
3. \( \zeta_k^{\text{opt}} \) depends on \( k \), \( \alpha^{\text{opt}} \) depends on \( m^* =: \min(m, n) \)
4. If \( k^* \) approx numerical rank \( A \), and right singular space is well-approximated \( \zeta_k^{\text{opt}} \approx \alpha^{\text{opt}} \) for \( \mathcal{K}_{k^*}(A^T A, A^T b) \)

Theorem (with assumptions on the approximation of the spectral space for RSVD follows from UPRE / WGCV convergence for FTSVD)

1. \( \alpha^{\text{opt}} \) for \( F^{\text{full}}(\alpha) \) can be estimated via RSVD
2. Minimizer of \( F^{\text{rsvd}}(\mu_k) \) is minimizer of \( F^{\text{full}}(\mu_k) \)
3. \( \mu_k^{\text{opt}} \) depends on \( k \), \( \alpha^{\text{opt}} \) depends on \( m^* =: \min(m, n) \)
4. For good rank \( k \) approximation \( \mu_k \approx \alpha_k \)
Regularization Estimation for UPRE / WGCV

Theorem ([RVA17]: with assumptions on the approximation of the spectral space for LSQR)

1. $\alpha^{\text{opt}}$ for $F^{\text{full}}(\alpha)$ can be estimated via LSQR
2. Minimizer of $F^{\text{proj}}(\zeta_k)$ is minimizer of $F^{\text{full}}(\zeta_k)$
3. $\zeta_k^{\text{opt}}$ depends on $k$, $\alpha^{\text{opt}}$ depends on $m^* =: \min(m, n)$
4. If $k^*$ approx numerical rank $A$, and right singular space is well-approximated $\zeta_{k^*}^{\text{opt}} \approx \alpha^{\text{opt}}$ for $K_k^*(A^T A, A^T b)$

Theorem (with assumptions on the approximation of the spectral space for RSVD follows from UPRE / WGCV convergence for FTSVD)

1. $\alpha^{\text{opt}}$ for $F^{\text{full}}(\alpha)$ can be estimated via RSVD
2. Minimizer of $F^{\text{rsvd}}(\mu_k)$ is minimizer of $F^{\text{full}}(\mu_k)$
3. $\mu_k^{\text{opt}}$ depends on $k$, $\alpha^{\text{opt}}$ depends on $m^* =: \min(m, n)$
4. For good rank $k$ approximation $\mu_k \approx \alpha_k$
Regularization Estimation for UPRE / WGCV

Theorem ([RVA17]: with assumptions on the approximation of the spectral space for LSQR)

1. $\alpha^{\text{opt}}$ for $F^{\text{full}}(\alpha)$ can be estimated via LSQR
2. Minimizer of $F^{\text{proj}}(\zeta_k)$ is minimizer of $F^{\text{full}}(\zeta_k)$
3. $\zeta_k^{\text{opt}}$ depends on $k$, $\alpha^{\text{opt}}$ depends on $m^* = \min(m, n)$
4. If $k^*$ approx numerical rank $A$, and right singular space is well-approximated $\zeta_{k^*}^{\text{opt}} \approx \alpha^{\text{opt}}$ for $K_{k^*}(A^T A, A^T b)$

Theorem (with assumptions on the approximation of the spectral space for RSVD follows from UPRE / WGCV convergence for FTSVD)

1. $\alpha^{\text{opt}}$ for $F^{\text{full}}(\alpha)$ can be estimated via RSVD
2. Minimizer of $F^{\text{rsvd}}(\mu_k)$ is minimizer of $F^{\text{full}}(\mu_k)$
3. $\mu_k^{\text{opt}}$ depends on $k$, $\alpha^{\text{opt}}$ depends on $m^* = \min(m, n)$
4. For good rank $k$ approximation $\mu_k \approx \alpha_k$
Contrasting the RSVĐ and LSQR

| **RSVD**  | RSVĐ with standard oversampling. \((p = k)\) |
| **RSVDQ** | RSVĐ with power iteration and \(q = 2\). \((p = k)\) |
| **LSQR**  | Standard LSQR |
| **LSQRO** | Oversample in the LSQR using \(p = k\) to find \(B_{k+p}\) and its SVD. Use relevant \(k\) components of the SVD as for the RSVĐ. |

**Aims**

1. Compare running times
2. Compare spectral approximation
3. Compare regularization estimation
**Contrasting the RSVDD and LSQR**

<table>
<thead>
<tr>
<th><strong>RSVD</strong></th>
<th>RSVD with standard oversampling. ((p = k))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RSVDQ</strong></td>
<td>RSVD with power iteration and (q = 2). ((p = k))</td>
</tr>
<tr>
<td><strong>LSQR</strong></td>
<td>Standard LSQR</td>
</tr>
<tr>
<td><strong>LSQRO</strong></td>
<td>Oversample in the LSQR using (p = k) to find (B_{k+p}) and its SVD. Use relevant (k) components of the SVD as for the RSVDD.</td>
</tr>
</tbody>
</table>

**Aims**

1. Compare running times
2. Compare spectral approximation
3. Compare regularization estimation
Contrasting the RSVD and LSQR

**RSVD**  RSVD with standard oversampling.  \( p = k \)

**RSVDQ**  RSVD with power iteration and  \( q = 2 \).  \( p = k \)

**LSQR**  Standard LSQR

**LSQRO**  Oversample in the LSQR using  \( p = k \) to find  \( B_{k+p} \) and its SVD. Use relevant  \( k \) components of the SVD as for the RSVD.

**Aims**
1. Compare running times
2. Compare spectral approximation
3. Compare regularization estimation
Contrasting the RSVD and LSQR

**RSVD** RSVD with standard oversampling. \( (p = k) \)

**RSVDQ** RSVD with power iteration and \( q = 2 \). \( (p = k) \)

**LSQR** Standard LSQR

**LSQRO** Oversample in the LSQR using \( p = k \) to find \( B_{k+p} \) and its SVD. Use relevant \( k \) components of the SVD as for the RSVD.

**Aims**
1. Compare running times
2. Compare spectral approximation
3. Compare regularization estimation
Contrasting the RSVDA and LSQR

RSVD  RSVD with standard oversampling. ($p = k$)
RSVDQ RSVD with power iteration and $q = 2$. ($p = k$)
LSQR  Standard LSQR
LSQRO Oversample in the LSQR using $p = k$ to find $B_{k+p}$ and its SVD. Use relevant $k$ components of the SVD as for the RSVDA.

Aims  1. Compare running times
  2. Compare spectral approximation
  3. Compare regularization estimation
Contrasting the RSVD and LSQR

**RSVD**  RSVD with standard oversampling. \((p = k)\)

**RSVDQ**  RSVD with power iteration and \(q = 2\). \((p = k)\)

**LSQR**  Standard LSQR

**LSQRO**  Oversample in the LSQR using \(p = k\) to find \(B_{k+p}\) and its SVD. Use relevant \(k\) components of the SVD as for the RSVD.

**Aims**

1. Compare running times
2. Compare spectral approximation
3. Compare regularization estimation
Contrasting the RSVD and LSQR

**RSVD**  RSVD with standard oversampling. \((p = k)\)

**RSVDQ**  RSVD with power iteration and \(q = 2\). \((p = k)\)

**LSQR**  Standard LSQR

**LSQRO**  Oversample in the LSQR using \(p = k\) to find \(B_{k+p}\) and its SVD. Use relevant \(k\) components of the SVD as for the RSVD.

**Aims**
1. Compare running times
2. Compare spectral approximation
3. Compare regularization estimation
Contrasting the RSVD and LSQR spectrum - impact of theory

Figure: RSVD

RSVD Singular Values

- True
- k=4
- k=12
- k=20
- k=28
- k=36
- k=44
- k=54
- k=60
Contrasting the RSVD and LSQR spectrum - impact of theory

Figure: LSQR

LSQR Singular Values

- True
- k=4
- k=12
- k=20
- k=28
- k=36
- k=44
- k=54
- k=60
Contrasting the RSVD and LSQR spectrum - impact of theory

Figure: LSQRO

LSQROver Singular Values

- True
- k=4
- k=12
- k=20
- k=28
- k=36
- k=44
- k=54
- k=60
Contrasting the RSVD and LSQR spectrum - impact of theory

Figure: RSVDQ

RSVDQ Singular Values
LSQR gives poor approximation of the singular space. LSQR with oversampling recovers accuracy comparable to RSVD.
Contrasting the RSVD and LSQR: singular space approximation

**Figure: run time**

LSQRO is expensive
Contrasting the RSVD and LSQR: convergence of regularization

The relative errors for phillips

Figure: No Regularization: Relative Error
Contrasting the RSVD and LSQR: convergence of regularization

Figure: Regularization: Relative Error

The relative errors for regularized phillips

- RSVD
- RSVDq
- LSQR
- LSQRO
Contrasting the RSVD and LSQR: convergence of regularization

Figure: Parameter Convergence

The regularization parameter \( \phi \) converges with \( k \) when singular space approximated well: RSVD, LSQRO, RSVDO

\( \alpha_k \) converges with \( k \) when singular space approximated well: RSVD, LSQRO, RSVDO.
Example Solutions for Phillips (Trivial)

The Regularized Solutions for phillips

Parameter $k$ increasing $[4, 12, 20, 28, 36, 44, 52, 60]$
Restoration of Grain noise level $\eta^2 = .0001$ : Restoretools

True and Contaminated

True Image

Blurred Noisy Image
Relative Errors decrease with TSVD approximation.
Summary Results: Image Restoration

Figure: Regularization Parameter

Regularization parameter converges as $k$ increases
Restored Regularized Solutions noise level $\eta^2 = .0001$ $k = 1200$ UPRE
Restored Regularized Solutions noise level $\eta^2 = .0001 \ k = 1200$ UPRE
Restored Regularized Solutions noise level $\eta^2 = .0001$ $k = 1200$ UPRE

Figure: LSQRO

LSQRREG
Restored Regularized Solutions noise level $\eta^2 = .0001$ $k = 1200$ UPRE

Figure: RSVDQ

LSQROREG
Optimal Solutions with RSVD and LSQR for Image Restoration Noise
5% oversampling 25%

Figure: RSVD UPRE $k = 2000$
Optimal Solutions with RSVD and LSQR for Image Restoration Noise
5% oversampling 25%

Figure: LSQR UPRE $k = 20$
Optimal Solutions with RSVD and LSQR for Image Restoration Noise
5% oversampling 25%

**Figure:** RSVD GCV $k = 2000$
Optimal Solutions with RSVD and LSQR for Image Restoration Noise
5% oversampling 25%

Figure: LSQR GCV \( k = 2000 \)
Table: The timings to restore the images illustrated for the Grain and Satellite Images

<table>
<thead>
<tr>
<th>Image</th>
<th>Grain</th>
<th>Satellite</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p=25</td>
<td>p=25</td>
</tr>
<tr>
<td>Oversampling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>UPRE</td>
<td>GCV</td>
</tr>
<tr>
<td>RSVD</td>
<td>54.602</td>
<td>55.646</td>
</tr>
<tr>
<td>LSQR</td>
<td>1.9909</td>
<td>1761.9</td>
</tr>
</tbody>
</table>

LSQR may require a smaller subspace \( k \) size

Notice solutions are still not good - blurred - lack of resolution
Iteratively Reweighted Regularization for Edge Preservation [LK83]

$$\|Ax - b\|^2 + \alpha^2 \|L(\ell)(x^{(\ell)} - x^{(\ell-1)})\|^2$$

**Minimum Support Stabilizer** Regularization operator $L(\ell)$.

$$(L(\ell))_{ii} = ((x_i^{(\ell-1)} - x_i^{(\ell-2)})^2 + \beta^2)^{-1/2} \quad \beta > 0$$

Parameter $\beta$ ensures $L(\ell)$ invertible

Invertibility use $(L(\ell))^{-1}$ as right preconditioner for $A$

$$(L(\ell))^{-1}_{ii} = ((x_i^{(\ell-1)} - x_i^{(\ell-2)})^2 + \beta^2)^{1/4} \quad \beta > 0$$

Initialization $L^{(0)} = I$, $x^{(0)} = x_0$. (might be 0)

Reduced System When $\beta = 0$ and $x_i^{(\ell-1)} = x_i^{(\ell-2)}$ remove column $i$, $\hat{A}$ is $AL^{-1}$ with columns removed.

Update Equation Solve $\hat{A}\hat{y} \approx R = b - Ax^{(\ell-1)}$. With correct indexing set $y_i = \hat{y}_i$ if updated, else $y_i = 0$.

$$x^{(\ell)} = x^{(\ell-1)} + y$$

**Cost of $L^{(\ell)}$ is minimal**
Iteratively Reweighted Regularization for Edge Preservation [LK83]

\[ \| Ax - b \| ^2 + \alpha ^2 \| L(\ell)(x(\ell) - x(\ell-1)) \| ^2 \]

**Minimum Support Stabilizer** Regularization operator \( L^{(\ell)} \).

\[(L^{(\ell)})_{ii} = (x_i^{(\ell-1)} - x_i^{(\ell-2)})^2 + \beta^2)^{-1/2} \quad \beta > 0\]

**Parameter** \( \beta \) ensures \( L^{(\ell)} \) invertible

**Invertibility** use \( (L^{(\ell)})^{-1} \) as right preconditioner for \( A \)

\[(L^{(\ell)})^{-1}_{ii} = ((x_i^{(\ell-1)} - x_i^{(\ell-2)})^2 + \beta^2)^{1/4} \quad \beta > 0\]

**Initialization** \( L^{(0)} = I, x^{(0)} = x_0 \). (might be 0)

**Reduced System** When \( \beta = 0 \) and \( x_i^{(\ell-1)} = x_i^{(\ell-2)} \) remove column \( i \), \( \hat{A} \) is \( AL^{-1} \) with columns removed.

**Update Equation** Solve \( \hat{A} \hat{y} \approx R = b - Ax^{(\ell-1)} \). With correct indexing set \( y_i = \hat{y}_i \) if updated, else \( y_i = 0 \).

\[ x^{(\ell)} = x^{(\ell-1)} + y \]

**Cost of** \( L^{(\ell)} \) **is minimal**
Iteratively Reweighted Regularization for Edge Preservation [LK83]

\[ \|Ax - b\|^2 + \alpha^2 \|L^{(\ell)}(x^{(\ell)} - x^{(\ell-1)})\|^2 \]

**Minimum Support Stabilizer** Regularization operator \( L^{(\ell)} \).

\[
(L^{(\ell)})_{ii} = ((x_i^{(\ell-1)} - x_i^{(\ell-2)})^2 + \beta^2)^{-1/2} \quad \beta > 0
\]

Parameter \( \beta \) ensures \( L^{(\ell)} \) invertible

Invertibility use \((L^{(\ell)})^{-1}\) as right preconditioner for \( A \)

\[
(L^{(\ell)})^{-1}_{ii} = ((x_i^{(\ell-1)} - x_i^{(\ell-2)})^2 + \beta^2)^{1/4} \quad \beta > 0
\]

Initialization \( L^{(0)} = I, \ x^{(0)} = x_0 \). (might be 0)

Reduced System When \( \beta = 0 \) and \( x_i^{(\ell-1)} = x_i^{(\ell-2)} \) remove column \( i \), \( \hat{A} \) is \( AL^{-1} \) with columns removed.

Update Equation Solve \( \hat{A}\hat{y} \approx R = b - Ax^{(\ell-1)} \). With correct indexing set \( y_i = \hat{y}_i \) if updated, else \( y_i = 0 \).

\[ x^{(\ell)} = x^{(\ell-1)} + y \]

**Cost of** \( L^{(\ell)} \) **is minimal**
Iteratively Reweighted Regularization for Edge Preservation [LK83]
\[ \|Ax - b\|^2 + \alpha^2 \|L^{(\ell)}(x^{(\ell)} - x^{(\ell-1)})\|^2 \]

**Minimum Support Stabilizer** Regularization operator \(L^{(\ell)}\).

\[ (L^{(\ell)})_{ii} = ((x_i^{(\ell-1)} - x_i^{(\ell-2)})^2 + \beta^2)^{-1/2} \quad \beta > 0 \]

Parameter \(\beta\) ensures \(L^{(\ell)}\) invertible

Invertibility use \((L^{(\ell)})^{-1}\) as right preconditioner for \(A\)

\[ (L^{(\ell)})^{-1}_{ii} = ((x_i^{(\ell-1)} - x_i^{(\ell-2)})^2 + \beta^2)^{1/4} \quad \beta > 0 \]

Initialization \(L^{(0)} = I, x^{(0)} = x_0\). (might be 0)

Reduced System When \(\beta = 0\) and \(x_i^{(\ell-1)} = x_i^{(\ell-2)}\) remove column \(i\), \(\hat{A}\) is \(AL^{-1}\) with columns removed.

Update Equation Solve \(\hat{A}\hat{y} \approx R = b - Ax^{(\ell-1)}\). With correct indexing set \(y_i = \hat{y}_i\) if updated, else \(y_i = 0\).

\[ x^{(\ell)} = x^{(\ell-1)} + y \]

**Cost of** \(L^{(\ell)}\) **is minimal**
Iteratively Reweighted Regularization for Edge Preservation [LK83]

\[ \|Ax - b\|^2 + \alpha^2 \|L^{(\ell)}(x^{(\ell)} - x^{(\ell-1)})\|^2 \]

**Minimum Support Stabilizer** Regularization operator \(L^{(\ell)}\).

\[
(L^{(\ell)})_{ii} = ((x_i^{(\ell-1)} - x_i^{(\ell-2)})^2 + \beta^2)^{-1/2} \quad \beta > 0
\]

Parameter \(\beta\) ensures \(L^{(\ell)}\) invertible

Invertibility use \((L^{(\ell)})^{-1}\) as right preconditioner for \(A\)

\[
(L^{(\ell)})^{-1}_{ii} = ((x_i^{(\ell-1)} - x_i^{(\ell-2)})^2 + \beta^2)^{1/4} \quad \beta > 0
\]

Initialization \(L^{(0)} = I, x^{(0)} = x_0.\) (might be 0)

Reduced System When \(\beta = 0\) and \(x_i^{(\ell-1)} = x_i^{(\ell-2)}\) remove column \(i\), \(\hat{A}\) is \(AL^{-1}\) with columns removed.

Update Equation Solve \(\hat{A}\hat{y} \approx R = b - Ax^{(\ell-1)}\). With correct indexing set \(y_i = \hat{y}_i\) if updated, else \(y_i = 0\).

\[
x^{(\ell)} = x^{(\ell-1)} + y
\]

Cost of \(L^{(\ell)}\) is minimal
Iteratively Reweighted Regularization for Edge Preservation [LK83]

\[ \| A x - b \|_2^2 + \alpha^2 \| L^{(\ell)} (x^{(\ell)} - x^{(\ell-1)}) \|_2^2 \]

**Minimum Support Stabilizer** Regularization operator \( L^{(\ell)} \).

\[
(L^{(\ell)})_{ii} = ((x^{(\ell-1)}_i - x^{(\ell-2)}_i)^2 + \beta^2)^{-1/2} \quad \beta > 0
\]

Parameter \( \beta \) ensures \( L^{(\ell)} \) invertible

Invertibility use \((L^{(\ell)})^{-1}\) as right preconditioner for \( A \)

\[
(L^{(\ell)})_{ii}^{-1} = ((x^{(\ell-1)}_i - x^{(\ell-2)}_i)^2 + \beta^2)^{1/4} \quad \beta > 0
\]

Initialization \( L^{(0)} = I, \; x^{(0)} = x_0 \). (might be 0)

Reduced System When \( \beta = 0 \) and \( x^{(\ell-1)}_i = x^{(\ell-2)}_i \) remove column \( i \), \( \hat{A} \) is \( AL^{-1} \) with columns removed.

Update Equation Solve \( \hat{A}\hat{y} \approx R = b - A x^{(\ell-1)} \). With correct indexing set \( y_i = \hat{y}_i \) if updated, else \( y_i = 0 \).

\[ x^{(\ell)} = x^{(\ell-1)} + y \]

**Cost of** \( L^{(\ell)} \) **is minimal**
Iteratively Reweighted Regularization for Edge Preservation [LK83]

\[ \|Ax - b\|^2 + \alpha^2 \|L^{(\ell)}(x^{(\ell)} - x^{(\ell-1)})\|^2 \]

**Minimum Support Stabilizer** Regularization operator \(L^{(\ell)}\).

\[
(L^{(\ell)})_{ii} = (\left((x_i^{(\ell-1)} - x_i^{(\ell-2)})^2 + \beta^2\right)^{-1/2} \quad \beta > 0
\]

Parameter \(\beta\) ensures \(L^{(\ell)}\) invertible

Invertibility use \((L^{(\ell)})^{-1}\) as right preconditioner for \(A\)

\[
(L^{(\ell)})^{-1}_{ii} = (\left((x_i^{(\ell-1)} - x_i^{(\ell-2)})^2 + \beta^2\right)^{1/4} \quad \beta > 0
\]

Initialization \(L^{(0)} = I, x^{(0)} = x_0\). (might be 0)

Reduced System When \(\beta = 0\) and \(x_i^{(\ell-1)} = x_i^{(\ell-2)}\) remove column \(i\), \(\hat{A}\) is \(AL^{-1}\) with columns removed.

Update Equation Solve \(\hat{A}\hat{y} \approx R = b - Ax^{(\ell-1)}\). With correct indexing set \(y_i = \hat{y}_i\) if updated, else \(y_i = 0\).

\[ x^{(\ell)} = x^{(\ell-1)} + y \]

**Cost of \(L^{(\ell)}\) is minimal**
Magnetic data $m = 5000, n = 75000 \beta^2 = 1e - 9, p = 10\%$

True \quad LSQR ($k = 5$) UPRE \quad RSVD ($k = 1000$) UPRE

LSQR slices - time 8.9566 seconds, $k = 5$

RSVD slices - time 1681.8 seconds, $k = 1000$
Magnetic data $m = 5000, n = 75000 \beta^2 = 1e - 9, p = 10\%$
Conclusions: RSVD - LSQR

**UPRE / WGCV** converges for the TSVD
**UPRE / WGCV** therefore converges for the RSVD
**UPRE / WGCV** converges for LSQR with oversampling

$$\zeta_k^{\text{opt}}, \alpha_k^{\text{opt}}, \mu_k^{\text{opt}}$$ related across levels for RSVD, RSVDQ and LSQRO

**Regularization** Find the optimal parameter for reduced subspace surrogate model and apply for larger number of terms.

**LSQR** Run with oversampling to avoid issues of semi-convergence but expensive

**RSVD or LSQR** Results suggest
- Advantages of the RSVD - speed!
- Disadvantage - not reflecting the full spectrum
UPRE / WGCV converges for the TSVD
UPRE / WGCV therefore converges for the RSVD
UPRE / WGCV converges for LSQR with oversampling
\( \zeta_{opt}^{k} \), \( \alpha_{opt}^{k} \), \( \mu_{opt}^{k} \) related across levels for RSVD, RSVDQ and LSQRO

Regularization Find the optimal parameter for reduced subspace surrogate model and apply for larger number of terms.

LSQR Run with oversampling to avoid issues of semi-convergence but expensive

RSVD or LSQR Results suggest

- Advantages of the RSVD - speed!
- Disadvantage - not reflecting the full spectrum
Conclusions: RSVD - LSQR

**UPRE / WGCV** converges for the TSVD
**UPRE / WGCV** therefore converges for the RSVD
**UPRE / WGCV** converges for LSQR with oversampling

ζₖ^{opt}, α^{opt}, μₖ^{opt} related across levels for RSVD, RSVDO and LSQRO

**Regularization** Find the optimal parameter for reduced subspace surrogate model and apply for larger number of terms.

**LSQR** Run with oversampling to avoid issues of semi-convergence but expensive

**RSVD or LSQR** Results suggest

- Advantages of the RSVD - speed!
- Disadvantage - not reflecting the full spectrum
Conclusions: RSVD - LSQR

UPRE / WGCV converges for the TSVD
UPRE / WGCV therefore converges for the RSVD
UPRE / WGCV converges for LSQR with oversampling

\( \zeta_k^{\text{opt}}, \alpha_k^{\text{opt}}, \mu_k^{\text{opt}} \) related across levels for RSVD, RSVDO and LSQRO

Regularization  Find the optimal parameter for reduced subspace surrogate model and apply for larger number of terms.

LSQR  Run with oversampling to avoid issues of semi-convergence but expensive

RSVD or LSQR  Results suggest
  ▶ Advantages of the RSVD - speed!
  ▶ Disadvantage - not reflecting the full spectrum
Conclusions: RSVD - LSQR

**UPRE / WGCV** converges for the TSVD

**UPRE / WGCV** therefore converges for the RSVD

**UPRE / WGCV** converges for LSQR with oversampling

\[ \zeta_k^{opt}, \alpha_k^{opt}, \mu_k^{opt} \] related across levels for RSVD, RSVDO and LSQRO

**Regularization** Find the optimal parameter for reduced subspace surrogate model and apply for a larger number of terms.

**LSQR** Run with oversampling to avoid issues of semi-convergence but expensive

**RSVD or LSQR** Results suggest

- Advantages of the RSVD - speed!
- Disadvantage - not reflecting the full spectrum
Conclusions: RSVD - LSQR

**UPRE / WGCV** converges for the TSVD
**UPRE / WGCV** therefore converges for the RSVD
**UPRE / WGCV** converges for LSQR with oversampling

\[ \zeta_{k}^{opt}, \alpha_{k}^{opt}, \mu_{k}^{opt} \] related across levels for RSVD, RSVDO and LSQRO

**Regularization** Find the optimal parameter for reduced subspace surrogate model and apply for larger number of terms.

- **LSQR** Run with oversampling to avoid issues of semi-convergence but expensive

- **RSVD or LSQR** Results suggest
  - Advantages of the RSVD - speed!
  - Disadvantage - not reflecting the full spectrum
Conclusions: RSVD - LSQR

**UPRE / WGCV** converges for the TSVD

**UPRE / WGCV** therefore converges for the RSVD

**UPRE / WGCV** converges for LSQR with oversampling

$\zeta_k^{opt}, \alpha_k^{opt}, \mu_k^{opt}$ related across levels for RSVD, RSVDQ and LSQRO

Regularization Find the optimal parameter for reduced subspace surrogate model and apply for larger number of terms.

**LSQR** Run with oversampling to avoid issues of semi-convergence but expensive

**RSVD or LSQR** Results suggest

- Advantages of the RSVD - speed!
- Disadvantage - not reflecting the full spectrum
**Low Rank**  Finding low rank approximation of model matrix is important (TRUNCATION)

**Benefits**  Low rank saves computational cost and memory

**Regularization**  Given low rank approximation estimate regularization parameter efficiently (FILTER)

**Cost**  While LSQR costs more per iteration, it converges faster in context of \( L_1 \) and is cheaper

**RSVD / LSQR**  Trade offs depend on speed by which singular values decrease

**RSVD**  Note algorithm is modified for \( m < n \).
Low Rank Finding low rank approximation of model matrix is important (TRUNCATION)

Benefits Low rank saves computational cost and memory

Regularization Given low rank approximation estimate regularization parameter efficiently (FILTER)

Cost While LSQR costs more per iteration, it converges faster in context of $L_1$ and is cheaper

RSVD / LSQR Trade offs depend on speed by which singular values decrease

RSVD Note algorithm is modified for $m < n$. 
Low Rank  Finding low rank approximation of model matrix is important (TRUNCATION)

Benefits  Low rank saves computational cost and memory

Regularization  Given low rank approximation estimate regularization parameter efficiently (FILTER)

Cost  While LSQR costs more per iteration, it converges faster in context of $L_1$ and is cheaper

RSVD / LSQR  Trade offs depend on speed by which singular values decrease

RSVD  Note algorithm is modified for $m < n$. 
Overview Conclusions

**Low Rank** Finding low rank approximation of model matrix is important (TRUNCATION)

**Benefits** Low rank saves computational cost and memory

**Regularization** Given low rank approximation estimate regularization parameter efficiently (FILTER)

**Cost** While LSQR costs more per iteration, it converges faster in context of $L_1$ and is cheaper

**RSVD / LSQR** Trade offs depend on speed by which singular values decrease

**RSVD** Note algorithm is modified for $m < n$. 
Low Rank  Finding low rank approximation of model matrix is important (TRUNCATION)

Benefits  Low rank saves computational cost and memory

Regularization  Given low rank approximation estimate regularization parameter efficiently (FILTER)

Cost  While LSQR costs more per iteration, it converges faster in context of $L_1$ and is cheaper

RSVD / LSQR  Trade offs depend on speed by which singular values decrease

RSVD  Note algorithm is modified for $m < n$. 
Overview

**Low Rank** Finding low rank approximation of model matrix is important (TRUNCATION)

**Benefits** Low rank saves computational cost and memory

**Regularization** Given low rank approximation estimate regularization parameter efficiently (FILTER)

**Cost** While LSQR costs more per iteration, it converges faster in context of $L_1$ and is cheaper

**RSVD / LSQR** Trade offs depend on speed by which singular values decrease

**RSVD** Note algorithm is modified for $m < n$. 

Some key references


Thank you.

Questions