Data Unfolding with Wiener-SVD Method

arXiv:1705.03568

W. Tang\textsuperscript{a,1} X. Li\textsuperscript{b,1} X. Qian\textsuperscript{a,2} H. Wei\textsuperscript{a} C. Zhang\textsuperscript{a}

\textsuperscript{a} Physics Department, Brookhaven National Laboratory, Upton, NY, USA
\textsuperscript{b} State University of New York at Stony Brook, Department of Physics and Astronomy, Stony Brook, NY, USA

xiaoyue.li@stonybrook.edu

DPF – FNAL, Aug. 2\textsuperscript{nd}, 2017
Outline

- The unfolding problem
- Wiener filter in digital signal processing
- SVD unfolding with Wiener filter
- Data unfolding example
  - Cross-section
  - Reactor neutrino flux
- Summary and recommendation
The data unfolding problem

- The problem:

\[ f_{\text{smear}}(x) = \int R(x|y)f_{\text{true}}(y) \, dy \]

- \( R(x|y) \): response function
- \( m_i = \sum_j R_{ij} s_j \)
- \( R_{ij} = P(\text{observed in bin } i \mid \text{true value in bin } j) \)
The data unfolding problem

- The problem:
  \[ f_{\text{smear}}(x) = \int R(x|y)f_{\text{true}}(y) \, dy \]
  - \( R(x|y) \): response function
  - \( m_i = \sum_j R_{ij} s_j \)
  - \( R_{ij} = P(\text{observed in bin } i \mid \text{true value in bin } j) \)
- Why not do \( s = R^{-1} \cdot m \)?
The data unfolding problem

- The problem:

$$f_{\text{smear}}(x) = \int R(x|y)f_{\text{true}}(y) \, dy$$

- $R(x|y)$: response function
- $m_i = \sum_j R_{ij} s_j$
- $R_{ij} = P(\text{observed in bin } i \mid \text{true value in bin } j)$
- Why not do $s = R^{-1} \cdot m$?
The data unfolding problem

- The problem:
  \[ f_{\text{smear}}(x) = \int R(x|y)f_{\text{true}}(y) \, dy \]
  - \( R(x|y) \): response function
  - \( m_i = \sum_j R_{ij}s_j \)
  - \( R_{ij} = P(\text{observed in bin } i \mid \text{true value in bin } j) \)

- Why not do \( s = R^{-1} \cdot m \)?
- Due to statistical fluctuation, the unfolded spectrum from direct inversion of response matrix bares no resemblance to the true spectrum.
The data unfolding problem

- The problem:
  \[
  f_{\text{smear}}(x) = \int R(x|y)f_{\text{true}}(y) \, dy
  \]
  - \( R(x|y) \): response function
  - \( m_i = \sum_j R_{ij}s_j \)
  - \( R_{ij} = P(\text{observed in bin } i \mid \text{true value in bin } j) \)

- Why not do \( s = R^{-1} \cdot m \)?
- Due to statistical fluctuation, the unfolded spectrum from direct inversion of response matrix bares no resemblance to the true spectrum.
- Introduce additional constraints.
  - E.g. trade bias for smoothness
  - Bayesian analysis
Wiener filter in digital signal processing (I)

- Deconvolution problem:
  - \( M(t') = \int_{-\infty}^{\infty} R(t, t') \cdot S(t) \, dt \)
  - True signal \( S(t) \), measured signal \( M(t') \), response function \( R(t, t') \equiv R(t - t') \)

![Diagram showing Wiener filter concepts](image)
Wiener filter in digital signal processing (1)

- Deconvolution problem:
  - \( M(t') = \int_{-\infty}^{\infty} R(t, t') \cdot S(t) \, dt \)
  - True signal \( S(t) \), measured signal \( M(t') \), response function \( R(t, t') \equiv R(t - t') \)
- Fourier transform:
  \( M(\omega) = R(\omega) \cdot S(\omega) \rightarrow S(\omega) = M(\omega)/R(\omega) \)
Wiener filter in digital signal processing (1)

- Deconvolution problem:
  - \( M(t') = \int_{-\infty}^{\infty} R(t, t') \cdot S(t) \, dt \)
  - True signal \( S(t) \), measured signal \( M(t') \), response function \( R(t, t') \equiv R(t - t') \)
- Fourier transform:
  \[
  M(\omega) = R(\omega) \cdot S(\omega) \rightarrow S(\omega) = M(\omega)/R(\omega)
  \]
  \( \rightarrow \) Inverse FFT \( S(\omega) \rightarrow S(t) \)
Wiener filter in digital signal processing (1)

- Deconvolution problem:
  - \[ M(t') = \int_{-\infty}^{\infty} R(t, t') \cdot S(t) \, dt \]
  - True signal \( S(t) \), measured signal \( M(t') \), response function \( R(t, t') \equiv R(t - t') \)

- Fourier transform:
  \[ M(\omega) = R(\omega) \cdot S(\omega) \rightarrow S(\omega) = \frac{M(\omega)}{R(\omega)} \]
  \[ \rightarrow \text{Inverse FFT} \quad S(\omega) \rightarrow S(t) \]

- The response function \( R(\omega) \) does not address noise contributions to the measured signal. Worse still, \( R(\omega) \) is generally smaller at higher frequencies due to the shaping features of electronics, resulting in amplification of noises.
Wiener filter in digital signal processing (2)

- To address the issue with noise:
  - $\hat{S}(\omega) = \frac{M(\omega)}{R(\omega)} \cdot F(\omega)$  
    Noise filter

- Wiener filter:
  - $F(\omega) = \frac{S^2(\omega)}{S^2(\omega) + [N^2(\omega)]}$  
    Noise

- The functional form of Wiener filter is obtained by minimizing
  $E[(F(\omega) \cdot M(\omega) - S(\omega))^2]$  
  $= E[(F(\omega) \cdot (S(\omega) + N(\omega)) - S(\omega))^2]$  

- **Wiener filter is designed to achieve the best signal/noise.**
Wiener filter in digital signal processing (3)
Reminder of SVD Method

- Minimize $\chi^2$:
  $$\chi^2(s) = (m - r \cdot s)^T \cdot Cov^{-1} \cdot (m - r \cdot s)$$

- Cholesky decomposition: $Cov^{-1} = Q^T \cdot Q$

- Pre-scaling: $M := Q \cdot m, R := Q \cdot r$

- ML estimator: $\hat{s} = (R^T R)^{-1} \cdot R^T \cdot M$

- Equivalently, the estimator can be written as
  $$\hat{s} = (R^T R)^{-1} \cdot R^T \cdot (R \cdot s_{\text{true}} + N)$$

where $N$, the “noise” coming from uncertainties, follows a normal distribution.
Reminder of SVD Method

- Minimize $\chi^2$:
  \[ \chi^2(s) = (m - r \cdot s)^T \cdot \text{Cov}^{-1} \cdot (m - r \cdot s) \]
- Cholesky decomposition: $\text{Cov}^{-1} = Q^T \cdot Q$
- Pre-scaling: $M := Q \cdot m$, $R := Q \cdot r$
- ML estimator: $\hat{s} = (R^T R)^{-1} \cdot R^T \cdot M$

Equivalently, the estimator can be written as:
\[ \hat{s} = (R^T R)^{-1} \cdot R^T \cdot (R \cdot s_{\text{true}} + N) \]

where $N$, the “noise” coming from uncertainties, follows a normal distribution.

Singular value decomposition (SVD)
\[ R = U \cdot D \cdot V^T \]

$U, V$: orthogonal
$D$: diagonal (diagonal elements ranked by magnitude)
Reminder of SVD Method

- Minimize $\chi^2$:
  \[ \chi^2(s) = (m - r \cdot s)^T \cdot Cov^{-1} \cdot (m - r \cdot s) \]
  - Cholesky decomposition: $Cov^{-1} = Q^T \cdot Q$
  - Pre-scaling: $M := Q \cdot m$, $R := Q \cdot r$
  - ML estimator: $\hat{s} = (R^T R)^{-1} \cdot R^T \cdot M$

- Equivalently, the estimator can be written as
  \[ \hat{s} = V \cdot D^{-1} \cdot U^T \cdot (R \cdot s_{true} + N) \]
  \[ = V \cdot D^{-1} \cdot (R_U \cdot s_{true} + N_U) = V \cdot D^{-1} \cdot M_U \]
  - Because $U$ is an orthogonal matrix, $N_U$ are still uncorrelated and follows a normal distribution.
  - $M_U = U^T \cdot M$ is the measurement in the effective frequency domain

Singular value decomposition (SVD) $R = U \cdot D \cdot V^T$
$U, V$: orthogonal
$D$: diagonal (diagonal elements ranked by magnitude)
Reminder of SVD Method

- Minimize $\chi^2$: 
  \[ \chi^2(s) = (m - r \cdot s)^T \cdot Cov^{-1} \cdot (m - r \cdot s) \]
  - Cholesky decomposition: $Cov^{-1} = Q^T \cdot Q$
  - Pre-scaling: $M := Q \cdot m$, $R := Q \cdot r$
  - ML estimator: $\hat{s} = (R^T R)^{-1} \cdot R^T \cdot M$

- Equivalently, the estimator can be written as
  \[
  \hat{s} = V \cdot D^{-1} \cdot U^T \cdot (R \cdot s_{true} + N) \]
  \[
  = V \cdot D^{-1} \cdot (R_U \cdot s_{true} + N_U) \]
  \[
  = V \cdot D^{-1} \cdot M_U
  \]
  - Because $U$ is an orthogonal matrix, $N_U$ are still uncorrelated and follows a normal distribution.
  - $M_U = U^T \cdot M$ is the measurement in the effective frequency domain

**Singular value decomposition (SVD)**

\[
R = U \cdot D \cdot V^T
\]

$U, V$: orthogonal
$D$: diagonal (diagonal elements ranked by magnitude)

- Recall in signal processing
  \[
  S(\omega) = [R(\omega)]^{-1} M(\omega)
  \]
Reminder of Regularization unfolding

- The estimator can be obtained by finding the maximum of
  \[ \phi(s) = \log L(s) + \tau \Sigma(s) \]
  Likelihood function
  Regularization function

- E.g. Tikhonov regularization
  \[ \Sigma^k(s(E)) = - \int \left( \frac{d^k s(E)}{d^k E} \right)^2 \, dE, k = 0, 1, 2, ... \]

- When \( k = 0 \), the unfolded result can be written as
  \[ \hat{s} = (R^T R + \tau I)^{-1} \cdot R^T \cdot M \]
  where \( I \) is an identity matrix and \( \tau \) is the regularization strength
Reminder of Regularization unfolding

- The estimator can be obtained by finding the maximum of
  \[ \phi(s) = \log L(s) + \tau \Sigma(s) \]

- E.g. Tikhonov regularization
  \[ \Sigma^k(s(E)) = - \int \left( \frac{d^k s(E)}{d^k E} \right)^2 dE, k = 0, 1, 2, ... \]

- When \( k = 0 \), the unfolded result can be written as
  \[ \hat{s} = (R^T R + \tau I)^{-1} \cdot R^T \cdot M \]

where \( I \) is an identity matrix and \( \tau \) is the regularization strength
Reminder of Regularization unfolding

- The estimator can be obtained by finding the maximum of
  \[ \phi(s) = \log L(s) + \tau \Sigma(s) \]
  - Likelihood function
  - Regularization function
  - Regularization strength

- E.g. Tikhonov regularization
  \[ \Sigma^k(s(E)) = - \int \left( \frac{d^k s(E)}{d^k E} \right)^2 dE, k = 0, 1, 2, ... \]
  - When \( k = 0 \), the unfolded result can be written as
    \[ \hat{s} = A \cdot (R^T R)^{-1} \cdot R^T \cdot M \]
    where \( A = V \cdot F \cdot V^T, F_{ij} = \frac{d_i^2}{d_i^2 + \tau} \delta_{ij} \) (\( d_i \) is the diagonal terms of \( D \))

\[ R = U \cdot D \cdot V^T \]
Reminder of Regularization unfolding

- The estimator can be obtained by finding the maximum of
  \[ \phi(s) = \log L(s) + \tau \Sigma(s) \]

- E.g. Tikhonov regularization

- When \( k = 0 \), the unfolded result can be written as
  \[ \hat{s} = A \cdot (R^T R)^{-1} \cdot R^T \cdot M \]

Where \( A = V \cdot F \cdot V^T \), \( F_{ij} = \frac{d_i^2}{d_i^2 + \tau} \delta_{ij} \) (\( d_i \) is the diagonal terms of \( D \))
Reminder of Regularization unfolding

- The estimator can be obtained by finding the maximum of
  \[ \phi(s) = \log L(s) + \tau \Sigma(s) \]

- E.g. Tikhonov regularization
  \[ \Sigma^k(s(E)) = -\int \left( \frac{\partial^k s(E)}{\partial E^k} \right)^2 dE, \quad k = 0, 1, 2, \ldots \]

- When \( k = 0 \), the unfolded result can be written as
  \[ \hat{s} = V \cdot F \cdot D^{-1} \cdot (R_U \cdot s_{\text{true}} + N_U), \quad \text{where} \quad F_{ij} = \frac{d_i^2}{d_i^2 + \tau} \delta_{ij} \]

\( N_U \) are suppressed by \( d_i / (d_i^2 + \tau) \), regardless of signal strength.
Recall: \( \hat{s} = V \cdot F \cdot D^{-1} \cdot (R_U \cdot s_{\text{true}} + N_U) \), \( F_{ij} = \frac{d_i^2}{d_i^2 + \tau} \delta_{ij} \)
Wiener-SVD

- Recall: \( \hat{s} = V \cdot W \cdot D^{-1} \cdot (R_U \cdot s_{\text{true}} + N_U) \), \( W \): Wiener filter
Wiener-SVD

- Recall: $\hat{s} = V \cdot W \cdot D^{-1} \cdot (R_U \cdot s_{\text{true}} + N_U)$, \(W\): Wiener filter

- In Wiener-SVD method, the functional form of \(W\) (replace \(F\) as filter) also factors in the signal expectation in the effective frequency domain

$$M_U := U^T \cdot \bar{M} = U^T \cdot R \cdot \bar{s} = D \cdot V^T \cdot \bar{s}$$

$$W_{ij} = \frac{M_{U_i}^2}{M_{U_i}^2 + N_{U_i}^2} \delta_{ij}$$
Wiener-SVD

- Recall: $\hat{s} = V \cdot W \cdot D^{-1} \cdot (R_U \cdot s_{\text{true}} + N_U)$, $W$: Wiener filter
- In Wiener-SVD method, the functional form of $W$ (replace $F$ as filter) also factors in the signal expectation in the effective frequency domain
  $\overline{M_U} := U^T \cdot \overline{M} = U^T \cdot R \cdot \overline{s} = D \cdot V^T \cdot \overline{s}$
- Signal: $\overline{M_{U_i}}^2 = d_i^2 \cdot (\sum_j V_{ij}^T \cdot \overline{s_j})^2$
- Noise: 1 (normal distribution)

$$W_{ij} = \frac{\overline{M_{U_i}}^2}{\overline{M_{U_i}}^2 + \overline{N_{U_i}}^2} \delta_{ij}$$
Wiener-SVD

- Recall: $\hat{s} = V \cdot W \cdot D^{-1} \cdot (R_U \cdot s_{\text{true}} + N_U)$, $W$: Wiener filter

- In Wiener-SVD method, the functional form of $W$ (replace $F$ as filter) also factors in the signal expectation in the effective frequency domain

\[ \overline{M_U} := U^T \cdot \overline{M} = U^T \cdot R \cdot \overline{s} = D \cdot V^T \cdot \overline{s} \]

- Signal: $\overline{M_{U_i}}^2 = d_i^2 \cdot (\sum_j V_{ij}^T \cdot \overline{s}_j)^2$

- Noise: 1 (normal distribution)

\[ W_{ij} = \frac{\overline{M_{U_i}}^2}{\overline{M_{U_i}}^2 + N_{U_i}} \delta_{ij} \]

\[ W_{ij} = \frac{d_i^2 \cdot (\sum_j V_{ij}^T \cdot \overline{s}_j)^2}{d_i^2 \cdot (\sum_j V_{ij}^T \cdot \overline{s}_j)^2 + 1} \delta_{ij} \]
In Tikhonov regularization, when \( k \neq 0 \), the solution is

\[
\hat{S} = \left( R^T R + \tau C^T C \right)^{-1} \cdot R^T \cdot \bar{M}
\]

where in the case of \( k = 2 \),

\[
C = \begin{pmatrix}
-1 & 1 & 0 & \cdots & 0 & 0 & 0 \\
1 & -2 & 1 & \cdots & 0 & 0 & 0 \\
0 & 1 & -2 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & -2 & 1 \\
0 & 0 & 0 & \cdots & 0 & -1 & 1
\end{pmatrix}
\]
Generalized Wiener-SVD

- In Tikhonov regularization, when $k \neq 0$, the solution is
  \[ \hat{s} = (R^T R + \tau \ C^T C)^{-1} \cdot R^T \cdot \bar{M} \]
  where in the case of $k = 2$,
  \[ C = \begin{pmatrix}
  -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\
  1 & -2 & 1 & \cdots & 0 & 0 & 0 \\
  0 & 1 & -2 & \cdots & 0 & 0 & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
  0 & 0 & 0 & \cdots & 1 & -2 & 1 \\
  0 & 0 & 0 & \cdots & 0 & -1 & 1
  \end{pmatrix} \]

- Similarly to Tikhonov regularization, we can add an additional matrix $C$ to the equation
  \[ \bar{M} = (R \cdot C^{-1}) \cdot (C \cdot \bar{s}) \]
Generalized Wiener-SVD

- In Tikhonov regularization, when $k \neq 0$, the solution is
  \[ \hat{s} = (R^T R + \tau C^T C)^{-1} \cdot R^T \cdot \bar{M} \]
  where in the case of $k = 2$,
  \[ C = \begin{pmatrix}
  -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\
  1 & -2 & 1 & \cdots & 0 & 0 & 0 \\
  0 & 1 & -2 & \cdots & 0 & 0 & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
  0 & 0 & 0 & \cdots & 1 & -2 & 1 \\
  0 & 0 & 0 & \cdots & 0 & -1 & 1
  \end{pmatrix} \]

- Similarly to Tikhonov regularization, we can add an additional matrix $C$ to the equation
  \[ \bar{M} = (R \cdot C^{-1}) \cdot (C \cdot \bar{s}) \]

Since the effective frequency domain is determined by the smearing matrix $R$, the inclusion of $C$ would alter the basis of the effective frequency domain.
Generalized Wiener-SVD

- In Tikhonov regularization, when $k \neq 0$, the solution is
  \[ \hat{s} = (R^T R + \tau C^T C)^{-1} \cdot R^T \cdot \bar{M} \]

  where in the case of $k = 2$,

  \[
  C = \begin{pmatrix}
  -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\
  1 & -2 & 1 & \cdots & 0 & 0 & 0 \\
  0 & 1 & -2 & \cdots & 0 & 0 & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
  0 & 0 & 0 & \cdots & 1 & -2 & 1 \\
  0 & 0 & 0 & \cdots & 0 & -1 & 1 \\
  \end{pmatrix}
  \]

- Similarly to Tikhonov regularization, we can add an additional matrix $C$ to the equation
  \[ \bar{M} = (R \cdot C^{-1}) \cdot (C \cdot \bar{s}) \]

\[ RC^{-1} = U_C \cdot D_C \cdot V_C^T \]
Generalized Wiener-SVD

- In Tikhonov regularization, when $k \neq 0$, the solution is
  \[ \hat{S} = (R^T R + \tau C^T C)^{-1} \cdot R^T \cdot \bar{M} \]

The Wiener-SVD solution becomes

\[ \hat{S} = C^{-1} \cdot V_C \cdot W_C \cdot V_C^T \cdot C \cdot (R^T R)^{-1} \cdot R^T \cdot \bar{M} \]

\[ (W_C)_{ij} = \frac{\sum_j (V_C^T)_{ij} \cdot (\sum_l c_{jl} \cdot \bar{s}_l)^2}{\sum_j (V_C^T)_{ij} \cdot (\sum_l c_{jl} \cdot \bar{s}_l)^2 + 1} \]

- Similarly to Tikhonov regularization, we can add an additional matrix $C$ to the equation
  \[ \bar{M} = (R \cdot C^{-1}) \cdot (C \cdot \bar{s}) \]

SVD

\[ RC^{-1} = U_C \cdot D_C \cdot V_C^T \]
Regularization interpretation of Wiener-SVD method

- The Wiener-SVD unfolding method is equivalent to regularization w.r.t. minimizing the signal to noise ratio in the effective frequency domain $(M_U)_i = \sum_j V_{ij} s_j$:

$$\phi(s) = \log L(s) + \frac{1}{2} \sum_i \log \frac{(M_U)_i^2}{N^2} = \log L(s) + \frac{1}{2} \sum_i \log \frac{(\sum_j V_{ij}^T s_j)^2}{1}$$
Application on neutrino cross-section measurements (1)

Toy experiment

Detector response

Statistics only uncertainties

Number of events

Energy

Reconstructed energy bin

True energy bin

Reconstructed energy bin

true spectrum

Asimov spectrum

measured spectrum

Statistics only uncertainties

Detector response

Number of events

Energy

Reconstructed energy bin

True energy bin

Reconstructed energy bin

true spectrum

Asimov spectrum

measured spectrum

Statistics only uncertainties

Detector response

Number of events

Energy

Reconstructed energy bin

True energy bin

Reconstructed energy bin

true spectrum

Asimov spectrum

measured spectrum

Statistics only uncertainties

Detector response

Number of events

Energy

Reconstructed energy bin

True energy bin

Reconstructed energy bin

true spectrum

Asimov spectrum

measured spectrum

Statistics only uncertainties

Detector response

Number of events

Energy

Reconstructed energy bin

True energy bin

Reconstructed energy bin

true spectrum

Asimov spectrum

measured spectrum

Statistics only uncertainties

Detector response

Number of events

Energy

Reconstructed energy bin

True energy bin

Reconstructed energy bin

true spectrum

Asimov spectrum

measured spectrum

Statistics only uncertainties

Detector response

Number of events

Energy

Reconstructed energy bin

True energy bin

Reconstructed energy bin

true spectrum

Asimov spectrum

measured spectrum

Statistics only uncertainties

Detector response

Number of events

Energy

Reconstructed energy bin

True energy bin

Reconstructed energy bin

true spectrum

Asimov spectrum

measured spectrum

Statistics only uncertainties

Detector response

Number of events

Energy

Reconstructed energy bin

True energy bin

Reconstructed energy bin

true spectrum

Asimov spectrum

measured spectrum

Statistics only uncertainties

Detector response

Number of events

Energy

Reconstructed energy bin

True energy bin

Reconstructed energy bin

true spectrum

Asimov spectrum

measured spectrum

Statistics only uncertainties

Detector response

Number of events

Energy

Reconstructed energy bin

True energy bin

Reconstructed energy bin

true spectrum

Asimov spectrum

measured spectrum

Statistics only uncertainties

Detector response

Number of events

Energy

Reconstructed energy bin

True energy bin

Reconstructed energy bin

true spectrum

Asimov spectrum

measured spectrum

Statistics only uncertainties

Detector response

Number of events

Energy

Reconstructed energy bin

True energy bin

Reconstructed energy bin

true spectrum

Asimov spectrum

measured spectrum

Statistics only uncertainties

Detector response
Application on neutrino cross-section measurements (2)

**Regularization**

\[
\hat{s} = V \cdot F \cdot D^{-1} \cdot M
\]

\[
F_{ij} = \frac{d_i^2}{d_i^2 + \tau} \delta_{ij}
\]

**Wiener-SVD**

\[
\hat{s} = V \cdot W \cdot D^{-1} \cdot M
\]

\[
W_{ii} = \frac{d_i^2 \cdot (\Sigma_j V_{ij}^T \cdot \bar{s}_j)^2}{d_i^2 \cdot (\Sigma_j V_{ij}^T \cdot \bar{s}_j)^2 + 1}
\]
Application on neutrino cross-section measurements (3)

Additional smearing matrix $A_c$

Unfolded covariance

\[ \hat{s} = A_c \cdot (R^T R)^{-1} \cdot R^T \cdot \overline{M} \]

Wiener-SVD w/ $C_2$

Tikhonov regularization w/ $C_2$
Application on neutrino cross-section measurements (4)

\[ \text{MSE} = \frac{1}{n} \left( \sigma^2 + b^2 \right) \]

\[ c_0 = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}, \]

\[ c_1 = \begin{bmatrix} -1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -1 & 1 \\ 0 & 0 & 0 & \ldots & 0 & -1 \end{bmatrix}, \]

\[ c_2 = \begin{bmatrix} -1 + \epsilon & 1 & 0 & \ldots & 0 & 0 \\ 1 & -2 + \epsilon & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -2 + \epsilon \\ 0 & 0 & 0 & \ldots & 1 & -1 + \epsilon \end{bmatrix}. \]
Application on reactor neutrino measurement (1)

Pre-scaled prompt energy distribution

\[ M := Q \cdot m \]

Covariance matrix includes both statistical and systematic uncertainties

Detector energy response matrix \( \mathbf{r} \) as reported by the Daya Bay collaboration

Pre-scaled detector energy response matrix \( R := Q \cdot \mathbf{r} \)

\[ Cov^{-1} = Q^T \cdot Q \]
Application on reactor neutrino measurement (2)

- In practice, the “true” energy spectrum is unknown; all we have are various models $\rightarrow$ take average

$$\overline{S^2} = M_U^2 = d_i^2 \cdot \frac{\sum_k \left( \sum_j V_{ij} \cdot s(k)_j \right)^2 \cdot e^{-\frac{x_k^2}{2}}}{\sum_k e^{-\frac{x_k^2}{2}}}$$

- The $\chi_k^2$ for model $k$ is calculated by comparing the predicted spectrum $m_k := r \cdot s_k$ and measurement $m$
Application on reactor neutrino measurement (3)

Regularization
Wiener-SVD
Wiener-SVD w/incorrect prediction

Using $\bar{s}$ to represent $s_{\text{true}}$

Using an incorrect $s$ to represent $s_{\text{true}}$

\[
\overline{S^2} = M_U^2 = d_i^2 \cdot \frac{\sum_k \left( \sum_j V_{ij}^T \cdot s(k)_j \right)^2 \cdot e^{-\frac{x_k^2}{2}}}{\sum_k e^{-\frac{x_k^2}{2}}}
\]
Discussions and recommendations (1)

- Pros & cons of Wiener-SVD
  - It does not require the evaluation of regularization strength.
  - It is model-dependent. However, the dependence can be mitigated by taking into account a number of different models, and by reporting the additional smearing matrix.

- Both the traditional regularization or Wiener-SVD are equivalent to applying a new smearing matrix, which suppresses the large oscillation (high variance) of the direct matrix inversion unfolded results, but also introduces bias.

\[
\hat{s} = A_C \cdot (R^T R)^{-1} \cdot R^T \cdot \bar{M}
\]

Where \( A_C = C^{-1} \cdot V_C \cdot W_C \cdot V_C^T \cdot C \)

\[
(W_C)_{ij} = \frac{d_i^2 \cdot \left( \Sigma_j (V_C^T)_{ij} \cdot (\Sigma_l C_{jl} \cdot \bar{s}_l) \right)^2}{d_i^2 \cdot \left( \Sigma_j (V_C^T)_{ij} \cdot (\Sigma_l C_{jl} \cdot \bar{s}_l) \right)^2 + 1}
\]
Discussions and recommendations (2)

- The Wiener filter should be constructed in ways which takes into account a range of prior expectations.
- $C_2$ matrix generally yields better results than $C_0$ and $C_1$.
- We recommend reporting this additional smearing matrix $A_C$ in the publication together with the unfolded results to enable a more direct comparison of expectations (e.g. from new theoretical calculations) with the unfolded results. In practice, the new smearing matrix should be applied to the theoretical calculation before comparing to the unfolded results.

arXiv:1705.03568

https://github.com/BNLIF/Wiener-SVD-Unfolding
Supplemental
Regularization interpretation of Wiener-SVD method

- The Wiener-SVD unfolding method is equivalent to regularization w.r.t. minimizing the signal to noise ratio in the effective frequency domain \((M_u)_i = \Sigma_j V^{T}_{ij} s_j\):

\[
\phi(s) = \log L(s) + \frac{1}{2} \sum_i \log \left( \frac{(M_u)_i^2}{N^2} \right) = \log L(s) + \frac{1}{2} \sum_i \log \left( \frac{\Sigma_j V^{T}_{ij} s_j}{1} \right)^2
\]

Minimizing Eq. (7) yields the following estimator

\[
\hat{s} = -X^{-1} \cdot Y \cdot M_u
\]

Where

\[
X_{ij} = \frac{\partial^2 \phi^2}{\partial s_i \partial s_j} = -(R^T R)_{ij} - \sum_k V_{ik} \cdot \frac{1}{(M_u)_k^2} \cdot V^T_{kj}
\]

\[
Y_{ij} = \frac{\partial^2 \phi^2}{\partial s_i \partial M_u_{ij}} = R^T_{ij}
\]

- With \(X\) and \(Y\) evaluated at the expectation of \(\bar{s}\) and \(\bar{M}\), one recovers Eq.(6)

\[
\hat{s} = V \cdot W \cdot V^T \cdot R^T \cdot \bar{M}
\]

With

\[
W_{ij} = \frac{d^2}{d^2_i + \frac{d^2}{M^2_i}} \delta_{ij} = \frac{d^2 \cdot (\Sigma_j V^{T}_{ij} \cdot \bar{s}_j)^2}{d^2_i \cdot (\Sigma_j V^{T}_{ij} \cdot \bar{s}_j)^2 + 1} \delta_{ij}
\]