Degree of correlation among influence parameters in total column ozone uncertainty budget determination through the double ratio method

O. El Gawhary[1]

[1]VSL Dutch Metrology Institute, The Netherlands

Symposium on uncertainty of ozone retrieval, Thessaloniki, April 2015

Outline

Double ratio method

Dobson and Brewer networks.
Total Column Ozone (TOC) measured at a set of wavelengths pairs. For instance, the *A* pair consists of the couple of wavelengths $\lambda_1 = 305.5$nm and $\lambda_2 = 325.4$nm.
Beer-Lambert law

$$I_\lambda = I_{0\lambda} \exp\left[-\alpha_\lambda \mu \Omega - \beta m \frac{P}{P_0} - \delta_\lambda \sec(Z)\right] \tag{1}$$

Double ratio method I

$$I_\lambda = I_{0\lambda} \exp\left[-\alpha_\lambda \mu \Omega - \beta m \frac{P}{P_0} - \delta_\lambda \sec(Z)\right] \qquad (2)$$

- I_λ is the direct normal spectra irradiance at λ
- $I_{0\lambda}$ is the extraterrestrial spectral irradiance at λ
- α_λ is the ozone absorption coefficient at λ
- μ is the ratio of actual and vertical paths of solar radiation through the ozone layer.
- Ω is the TOC
- β_λ is the Rayleigh scattering coefficient at λ
- m is the airmass corresponding to solar zenith.
- P is the atmospheric pressure at the measurement station
- P_0 is the mean sea pressure
- δ_λ is the scattering coefficient (optical depth) of aerosol at wavelength λ.
- Z is the solar zenith angle

Double ratio method II

If the spectral irradiance is measured at one pair of wavelengths, then one can, in principle, obtain a value of the O_3 by inverting

$$I_\lambda = I_{0\lambda} \exp\left[-\alpha_\lambda \mu \Omega - \beta m \frac{P}{P_0} - \delta_\lambda \sec(Z)\right] \quad (3)$$

$$\Omega = \frac{N - [(\beta - \beta')mP/P_0] - (\delta - \delta')\sec(Z)}{(\alpha - \alpha')\mu} \quad (4)$$

where $N = \log I/I_0' - \log I/I'$

Double ratio method III

If the measurements at two distinct couples of wavelengths are combined together one gets

$$\Omega = \frac{(N_1 - N_2) - [(\beta - \beta')_1 - (\beta - \beta')_2]\, mP/P_0 - [(\delta - \delta')_1 - (\delta - \delta')_2]\, \sec(Z)}{[(\alpha - \alpha')_1 - (\alpha - \alpha')]\, \mu} \tag{5}$$

It is generally assumed that $(\delta - \delta')_1 - (\delta - \delta')_2 \simeq 0$

Double ratio method IV

The process of determining the total column ozone can be seen as a method where one tries to match the measured quantity

$$y^{(m)} = (\log I/I_0' - \log I/I')_1 - (\log I/I_0' - \log I/I')_2 \quad (6)$$

and the model

$$y = (\Delta\alpha_1 - \Delta\alpha_2)\,\mu\Omega + (\Delta\beta_1 - \Delta\beta_2)\,mP/P_0 + \Delta\delta\,\sec(Z) \quad (7)$$

where
$\Delta\alpha = \alpha - \alpha'$, $\Delta\alpha = \beta - \beta'$ and $\Delta\delta = (\delta - \delta')_1 - (\delta - \delta')_2$.

Jacobian matrix

$\Delta\alpha_1$, $\Delta\alpha_2$, $\Delta\beta_1$,Ω, $\Delta\beta_2$, P, $\Delta\delta$, Z (μ and m are function of Z).
We build the Jacobian matrix

$$[J_{jk}] = \left[\frac{\partial y}{\partial a_j} \cdot \frac{\partial y}{\partial a_k}\right] \tag{8}$$

$a_j \Rightarrow$ parameter, with $j = 1, \ldots 8$, $a_1 = \Delta\alpha_1$, $a_2 = \Delta\alpha_2$, and so on.
From $[J_{jk}]$ one can compute the covariance matrix
$[C_{jk}] = [J_{jk}]^{-1}$.
Degree of cross correlation matrix $[\rho_{jk}]$

$$[\rho_{jk}] = \left[\frac{C_{jk}}{\sqrt{C_{jj}}\sqrt{C_{kk}}}\right] \tag{9}$$

Jacobian matrix

But $\left[J_{jk}\right]$ must be invertible or not ill-conditioned.
Unfortunately, this is exactly our case.

$a_1 = \Delta\alpha_2$, $a_2 = \Delta\alpha_1$, $a_3 = \Omega$, $a_4 = \Delta\beta_2$, $a_5 = \Delta\beta_1$, $a_6 = P$, $a_7 = \Delta\delta$ and $a_8 = Z$

$$J_{11} = (\mu\Omega)^2$$
$$J_{12} = -(\mu\Omega)^2$$
$$J_{13} = \mu^2\Omega(\Delta\alpha_2 - \Delta\alpha_1)$$
$$J_{14} = \mu\Omega m \frac{P}{P_0}$$
$$J_{15} = -\mu\Omega m \frac{P}{P_0}$$
$$J_{16} = \mu\Omega \frac{(\Delta\beta_2 - \Delta\beta_1)}{P_0}$$
$$J_{17} = \mu\Omega \sec(Z)$$
$$J_{18} = \mu\Omega \left[(\Delta\alpha_2 - \Delta\alpha_1)\,\Omega\frac{d\mu}{dZ} + (\Delta\beta_2 - \Delta\beta_1)\,\frac{P}{P_0}\frac{dm}{dZ} + \Delta\delta\frac{d\sec(Z)}{dZ}\right]$$

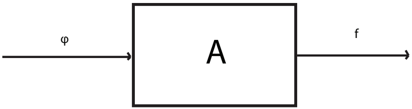
Example. Ozone observatory center at Boulder, Colorado.

1. $\Delta\alpha_2 = 1.806 \left[(atm - cm)^{-1}\right]$
2. $\Delta\alpha_1 = 0.374 \left[(atm - cm)^{-1}\right]$
3. $\Delta\beta_2 = 0.114 \left[(atm)^{-1}\right]$
4. $\Delta\beta_1 = 0.104 \left[(atm)^{-1}\right]$
5. $Z = 63.448\,[degrees]$
6. $\Omega = 0.355\,[atm - cm]$
7. $\mu = 2.211807$
8. $P_0 = 1.0193295 \cdot 10^3 [mbars]$
9. $P = 0.840997499 \cdot 10^3 [mbars]$
10. $m = 2.2289096$
11. $\Delta\delta = 0;$

J_{jk} with a condition number $4.0523925 \cdot 10^{32}$. This is in fact a clear indication of the close-to-singularity condition suffered by the Jacobian matrix J_{jk} .

Well and ill-posed problems

Let us consider a linear system

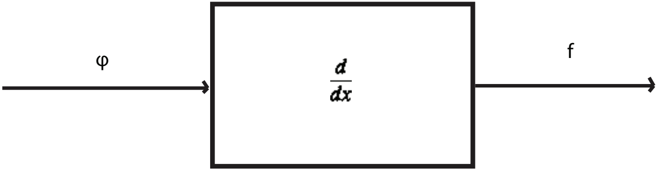


$$A\varphi = f$$

A is the direct operator.
φ is the input.
f is the output.

Well-posed problems: an example I

$$A = \frac{d}{dx}$$



If

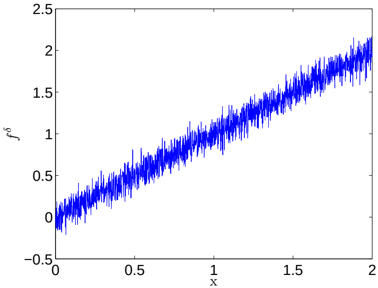
$$\varphi = \frac{x^2}{2} \text{ then} \Rightarrow A\varphi = f = x$$

Well-posed problems: an example II

Classical inverse problem:
From a measured $f^{\delta}(x) = f(x) + noise$, and known A, obtain φ.

$$\text{Since } A\varphi = f \text{ then } \varphi^{\delta} = A^{-1}f^{\delta}. \ (AA^{-1} = I)$$

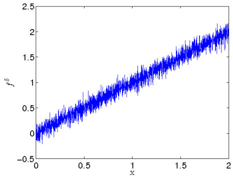
Example: $f(x) = x$, then $f^{\delta}(x)$ is (5% noise)



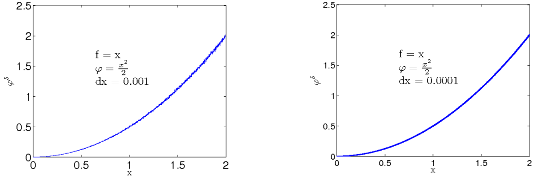
Well-posed problems: an example III

A^{-1} is an integration $\Rightarrow \varphi^{\delta}(x) = \int_0^x [\xi + noise] d\xi$

from measurements (5% noise)

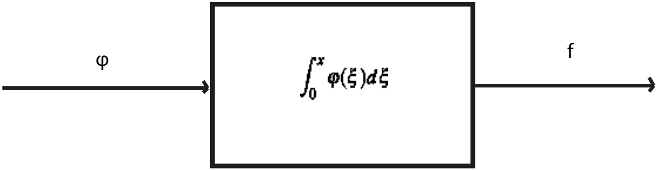


after applying A-1



Ill-posed problems: an example I

$A = \int_0^x (\cdot) d\xi$



If

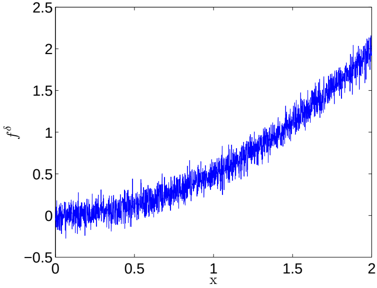
$$\varphi = x \text{ then} \Rightarrow A\varphi = f = \frac{x^2}{2}.$$

Ill-posed problems: an example II

From a measured $f^{\delta}(x) = f(x) + noise$, and known A, obtain φ.

$$\text{Since } A\varphi = f \text{ then } \varphi^{\delta} = A^{-1}f^{\delta}.$$

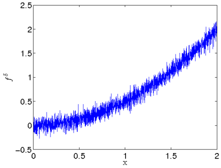
Example: $f(x) = \frac{x^2}{2}$, then $f^{\delta}(x)$ is (5% noise)



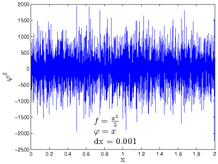
Ill-posed problems: an example III

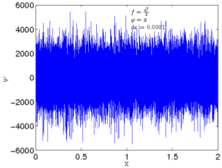
A^{-1} is an differentiation $\Rightarrow \varphi^{\delta}(x) = \frac{d}{dx}[\frac{x^2}{2} + noise]$

from measurements (5% noise)



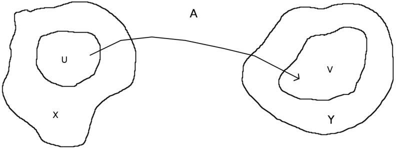
after applying A-1





How to cure a problem which is ill-posed

Hadamard's concept of well-posedness



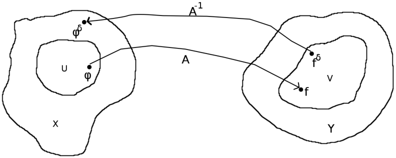
$A : U \subset X \longrightarrow V \subset Y$

The equation $A\varphi = f$ is well-posed if

1. A is bijective
2. A^{-1} is continuous (discontinuity leads to instability with respect to noise)

Regularization methods

A *regularization method* is a method to construct a stable (with respect to noise) approximate solution of an ill-posed problem of interest.



Aim: to compute an approximation of φ when a noisy version of f is known.

$$\| f^{\delta} - f \| \leq \delta, \ \delta \text{ is the noise level.}$$

Approximation for the inverse operator

We have the *erroneous* data f^δ available and we want to obtain a good approximation φ^δ of φ.
The problem usually resides in the form of A^{-1}.
Goal:

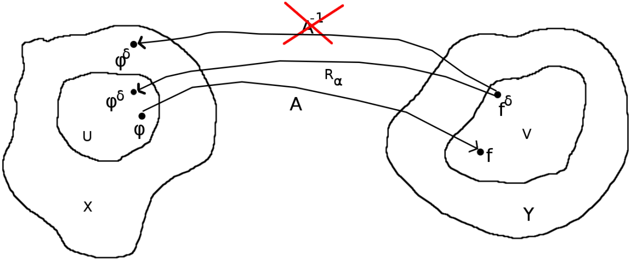
To find an approximation of the unbounded inverse operator A^{-1} such that φ^δ depends continuously on f^δ

Definition:
The family of bounded operators R_α such that

$$\lim_{\alpha\to 0} R_\alpha A\varphi = \varphi$$

is called a regularization scheme for the operator A

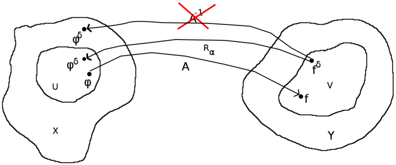
Approximation for the inverse operator



One approximates the solution φ by $\varphi_\alpha^\delta = R_\alpha f^\delta$

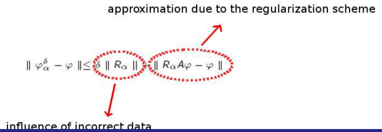
Problem: R_α is NOT the inverse of A for $\alpha \neq 0$!

Stability vs. accuracy



$$\varphi_\alpha^\delta - \varphi = R_\alpha f^\delta - R_\alpha f + R_\alpha A\varphi - \varphi$$

$$\text{total error} = \| \varphi_\alpha^\delta - \varphi \| \leq \delta \| R_\alpha \| + \| R_\alpha A\varphi - \varphi \|$$



A look at the origin of ill-posedness

We have to recall few facts on the spectral decomposition of an operator.
$A : X \rightarrow Y$ is a compact linear operator.
$A^a : X \rightarrow Y$ is the adjoint operator of A with property:

$$(A\varphi, \psi) = (\varphi, A^a\psi)$$

$\forall\varphi \in X$ and $\psi \in Y$.
When $A = A^a$, A is said to be hermitian.
The operator $A^aA : X \rightarrow X$ is hermitian.

$$A^aA\psi_n = \mu_n^2\psi_n$$

$\mu_1 \geq \mu_2 \geq \mu_3 \geq \dots$ are called singular values of A.
(Note that $\mu_n \rightarrow 0$, when $n \rightarrow \infty$)

A look at the origin of ill-posedness: singular value decomposition

Given the sequence of singular values μ_n then there exist two sequences

$$\{\varphi_n\} \text{ in } X \text{ and } \{g_n\} \text{ in } Y$$

such that

$$A\varphi_n = \mu_n g_n$$
$$A^a g_n = \mu_n \varphi_n$$

It can be proven that any φ in X can be expanded as

$$\varphi = \sum_{n=1}^{\infty} (\varphi, \varphi_n)\varphi_n$$

Finally, given an operator $A : X \to Y$, with a singular system (μ_n, φ_n, g_n), then the problem $A\varphi = f$ can be solved for φ and gives

$$\varphi = \sum_{n=1}^{\infty} \frac{1}{\mu_n} (f, g_n)\varphi_n$$

(Picard's theorem)

Consequence of Picard's theorem: mild and severe ill-posedness

Picard's theorem shows the ill-posed nature of an equation.

$$A\varphi = f$$

Let us suppose we perturb f

$$f^{\delta} = f + \delta\, g_n$$

The expansion $\varphi = \sum_{n=1}^{\infty} \frac{1}{\mu_n}(f, g_n)\varphi_n$ implies

$$\varphi^{\delta} = \varphi + \delta \frac{\varphi_n}{\mu_n}$$

It follows that the ratio

$$\frac{\|\varphi^{\delta} - \varphi\|}{\|f^{\delta} - f\|} = \frac{1}{\mu_n}$$

can become arbitrarily large since μ_n tend to zero.

Small variations in the measurements lead to large variations in the reconstruction (instability)

Mild and severe ill-posedness and regularization schemes

The equation

$$A\varphi = f$$

is said to be

1. mildly ill-posed when μ_n decay slowly to zero
2. severely ill-posed when μ_n decay rapidly to zero

Regularization schemes

Two of the most used regularization schemes are

- Tikhonov regularization
- Spectral cut-off

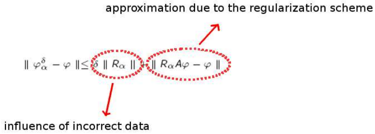
These are methods to build the approximated inverse operator R_α.

Tikhonov regularization

Under some hypotheses the following regularization schemes applies

$$R_\alpha = (\alpha I + A^a A)^{-1} A^a$$

From what we said before



there is an optimum α (often found by trial and error)

Spectral cut-off

In the spectral cut-off one simply replaces the complete expansion

$$\varphi = \sum_{n=1}^{\infty} (\varphi, \varphi_n) \varphi_n$$

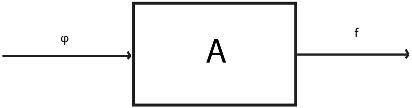
by the finite sum

$$R_N f := \sum_{n=1}^{N} \frac{1}{\mu_n} (f, g_n) \varphi_n$$

In optics, this method leads to the diffraction limit.

An example of Tikhonov regularization

Let us assume we have the following linear system:

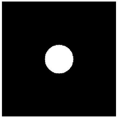


where

$$f(\mathbf{x}) = \int \varphi(\xi) h(\mathbf{x} - \xi) d\xi$$

The operator A is then a convolution operator. If we set


φ


h

An example of Tikhonov regularization II

Now let us assume we know the structure of the operator A, we measure the output f and we want to retrieve the incident field φ (input).
In a Fourier domain $f(\boldsymbol{x}) = \int \varphi(\xi)h(\boldsymbol{x} - \xi)d\xi$ becomes

$$\tilde{f}(\boldsymbol{p}) = \tilde{\varphi}(\boldsymbol{p})\tilde{h}(\boldsymbol{p})$$

Inverse problem: from the measured f try to get the unknown φ.
We can try the following inversion scheme

$$\tilde{\varphi}(\boldsymbol{p}) = \frac{\tilde{f}(\boldsymbol{p})}{\tilde{h}(\boldsymbol{p})}$$

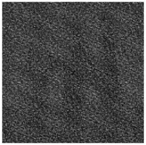
An example of Tikhonov regularization III

We actually measure the noisy version f^δ

$$\tilde{\varphi}(\boldsymbol{p}) = \frac{\tilde{f^\delta}(\boldsymbol{p})}{\tilde{h}(\boldsymbol{p})}$$

We get

φ



φ^δ

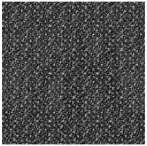
Clearly the reconstruction through the formal inversion is completely ill-posed.

An example of Tikhonov regularization IV

Let us try to regularize the solution by a Tikhonov scheme.

$$\tilde{\varphi}^{\delta}(\boldsymbol{p}) = \tilde{f}^{\delta}(\boldsymbol{p}) \frac{(\tilde{h}(\boldsymbol{p}))^{\star}}{[\alpha + \tilde{h}(\boldsymbol{p})\tilde{h}(\boldsymbol{p})^{\star}]}$$

φ



φ^{δ}

$\alpha = 0.000001$ is still too small to get a stable solution.

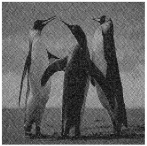
An example of Tikhonov regularization V

The right regularization parameter α depends on the noise level.

Let us choose now $\alpha = 0.001$.

$$\tilde{\varphi}^{\delta}(\boldsymbol{p}) = \tilde{f}^{\delta}(\boldsymbol{p}) \frac{(\tilde{h}(\boldsymbol{p}))^{\star}}{[\alpha + \tilde{h}(\boldsymbol{p})\tilde{h}(\boldsymbol{p})^{\star}]}$$

φ



φ^{δ}

We see that the effect of noise has been notably reduced.

An example of Tikhonov regularization V

Let us try another α.
Let us choose now $\alpha = 0.1$.

$$\tilde{\varphi}^{\delta}(\boldsymbol{p}) = \tilde{f}^{\delta}(\boldsymbol{p})\frac{(\tilde{h}(\boldsymbol{p}))^{\star}}{[\alpha+\tilde{h}(\boldsymbol{p})\tilde{h}(\boldsymbol{p})^{\star}]}$$

φ

φ^{δ}

An example of Tikhonov regularization VI

Let us choose now $\alpha = 10$.

$$\tilde{\varphi}^{\delta}(\boldsymbol{p}) = \tilde{f}^{\delta}(\boldsymbol{p}) \frac{(\tilde{h}(\boldsymbol{p}))^{\star}}{[\alpha + \tilde{h}(\boldsymbol{p})\tilde{h}(\boldsymbol{p})^{\star}]}$$

φ

φ^{δ}

Regularization stabilizes the solution but does not provide a perfect inversion.

An example of Tikhonov regularization VI

If we try to regularize too much.....
Let us choose now $\alpha = 1000$.

$$\tilde{\varphi}^{\delta}(\boldsymbol{p}) = \tilde{f}^{\delta}(\boldsymbol{p}) \frac{(\tilde{h}(\boldsymbol{p}))^{\star}}{[\alpha + \tilde{h}(\boldsymbol{p})\tilde{h}(\boldsymbol{p})^{\star}]}$$

φ

φ^{δ}

Singular Value Decomposition

Coming back to our original problem... For any matrix is however possible to find three matrices U, W, V (that in this specific case are all 8×8 square matrices) such that

$$J = U \cdot W \cdot V^T \tag{10}$$

where U and V are orthogonal matrices so that their inverse are equal to their transposes.
W is a diagonal matrix with its diagonal (all positive) elements being the singular values of the original matrix J.
Written in this way, the inverse J^{-1} would take the form

$$J^{-1} = V \cdot \left[diag(1/w_j)\right] \cdot V^T \tag{11}$$

Jacobian matrix

The covariance matrix can be expressed as

$$\left[C_{jk}\right] = \sum_{i=1}^{8}\left(\frac{V_{ji}}{w_i}\right)^2$$

$\left[\rho_{jk}\right]$ becomes, after applying a spectral cut-off to our original Jacobian $\left[J_{jk}\right]$

$$\begin{bmatrix}
1 & -1 & -0.81729 & -0.93193 & 0.93193 & -0.93193 & 0.74638 & -0.93193 \\
-1 & 1 & 0.81729 & 0.93193 & -0.93193 & 0.93193 & -0.74638 & 0.93193 \\
-0.81729 & 0.81729 & 1 & 0.90002 & -0.90002 & 0.90002 & -0.35610 & 0.90002 \\
-0.93193 & 0.931932 & 0.90002 & 1 & -1 & 0.99999 & -0.45424 & 0.99999 \\
0.931932 & -0.931932 & -0.90002 & -1 & 1 & -0.99999 & 0.45424 & -1 \\
-0.93193 & 0.93193 & 0.90002 & 0.99999 & -0.99999 & 1 & -0.45424 & 1 \\
0.74638 & -0.74638 & -0.35610 & -0.45424 & 0.45424 & -0.45424 & 1 & -0.45424 \\
-0.93193 & 0.93193 & 0.90002 & 0.99999 & -1 & 1 & -0.45424 & 1
\end{bmatrix}$$

$a_1 = \Delta\alpha_2$, $a_2 = \Delta\alpha_1$, $a_3 = \Omega$, $a_4 = \Delta\beta_2$, $a_5 = \Delta\beta_1$, $a_6 = P$, $a_7 = \Delta\delta$ and $a_8 = Z$

Conclusions

- We have discussed a possible way to estimate the degree of correlation for the influence parameters in TOC uncertainty calculations coming from the measurement equation.
- We have to include the correlations coming from the calibration. How to combine them?
- Application to series of data. Please, share you TOC data.