SVD and The Naive Solution

Recall that we define the SVD of the $N \times N$ matrix $\mathbf{A}$ to be

$$\mathbf{A} = \mathbf{U}\,\boldsymbol{\Sigma}\,\mathbf{V}^T,$$

where $\mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = \mathbf{I}_N$, and $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_N)$ with

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_N \geq 0.$$

All the singular values decay gradually to zero, and the condition number $\mathrm{cond}(\mathbf{A}) = \sigma_1/\sigma_N$ is very large (practically infinite).

Given $\mathbf{b} = \mathbf{b}_{\mathrm{exact}} + \mathbf{e}$ (pure data plus noise), the naive solution

$$\mathbf{x}_{\mathrm{naive}} = \mathbf{A}^{-1}\mathbf{b} = \mathbf{A}^{-1}\mathbf{b}_{\mathrm{exact}} + \mathbf{A}^{-1}\mathbf{e} = \mathbf{x} + \sum_{i=1}^{N} \frac{\mathbf{u}_i^T\mathbf{e}}{\sigma_i}\,\mathbf{v}_i$$

is completely dominated by the inverted noise component $\mathbf{A}^{-1}\mathbf{e}$.

Truncated SVD (TSVD)

A simple approach to noise reduction in the reconstruction:

$$\boxed{\text{Discard all SVD components that are dominated by noise.}}$$

As we shall soon see, these components are typically the ones for indices i above a certain truncation parameter k.

This leads to the *TSVD solution*

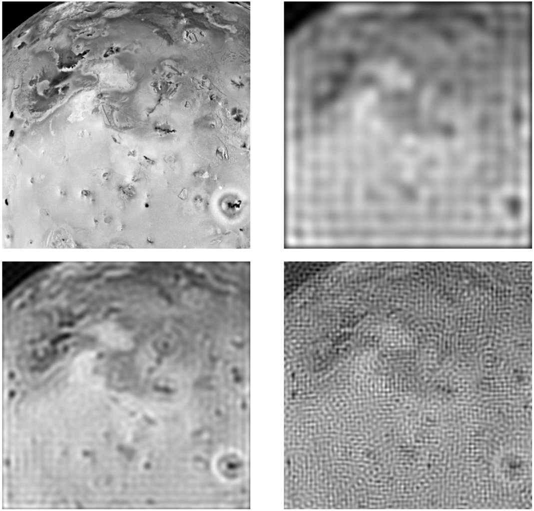
$$\mathbf{x}_k = \sum_{i=1}^{k} \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i, \qquad k < N.$$

Works well – in spite of its simplicity (the explanation follows).

The next overhead shows an example (Io image) with

- atmospheric turbulence blur and
- 5% Gaussian white noise.

Examples of TSVD Solutions: $k = 568,\ 2813,\ 7243$



Spectral Filtering – General Formulation

TSVD is an example of the general class of methods that are called *spectral filtering* methods, and which have the form

$$\mathbf{x}_{\text{filt}} = \sum_{i=1}^{N} \phi_i \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i \,,$$

- the quantities ϕ_i are the *filter factors*;
- they are chosen such that $\phi_i \approx 1$ for large singular values, and $\phi_i \approx 0$ for small singular values;
- different regularization algorithms involve different choices of these filter factors ($\rightarrow$ Chapter 6).

The SVD formulation is primarily used for defining the spectral filtering methods. Computational algorithms should only use the SVD – or the spectral decomposition – if it can be computed fast!

Incorporating Boundary Conditions

Recall (from Chapter 4) that to incorporate boundary conditions, we should really work with the deblurring model

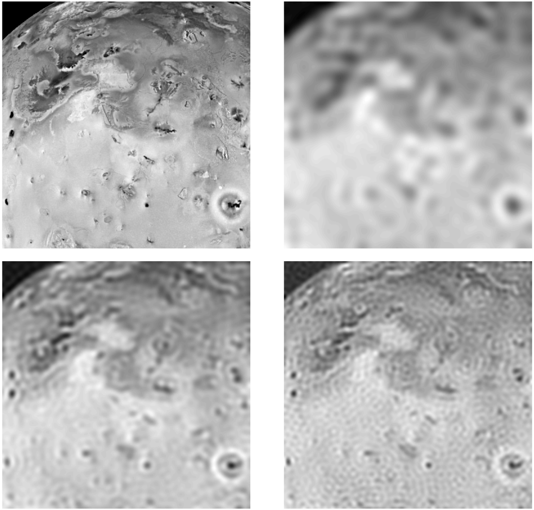
$$\mathbf{A}\,\mathbf{x} = \mathbf{b}, \qquad \text{with} \qquad \mathbf{A} = \mathbf{A}_0 + \mathbf{A}_{\mathrm{BC}},$$

where

- $\mathbf{A}_0$ is the structured BTTB matrix resulting from zero boundary conditions, and
- $\mathbf{A}_{\mathrm{BC}}$ is a correction term that incorporates specific boundary conditions into the model.
- The matrix $\mathbf{A}_{\mathrm{BC}}$ is also structured, and its form depends on the type of boundary condition.

Use the SVD of the corrected matrix $\mathbf{A}$ (not the SVD of $\mathbf{A}_0$).

TSVD and Reflexive Boundary Conditions



Very Important Points, So Far

1. Our fundamental decomposition is either the singular value decomposition or the spectral decomposition.
2. Spectral filtering amounts to filtering each of the components of the solution in the spectral basis, in such a way that the influence from the noise in the blurred image is damped.
3. In order to obtain a high quality deblurred image, we must choose the boundary conditions appropriately.
4. Reconstructions based on $\mathbf{A} = \mathbf{A}_0$ correspond to zero boundary conditions.

SVD Analysis

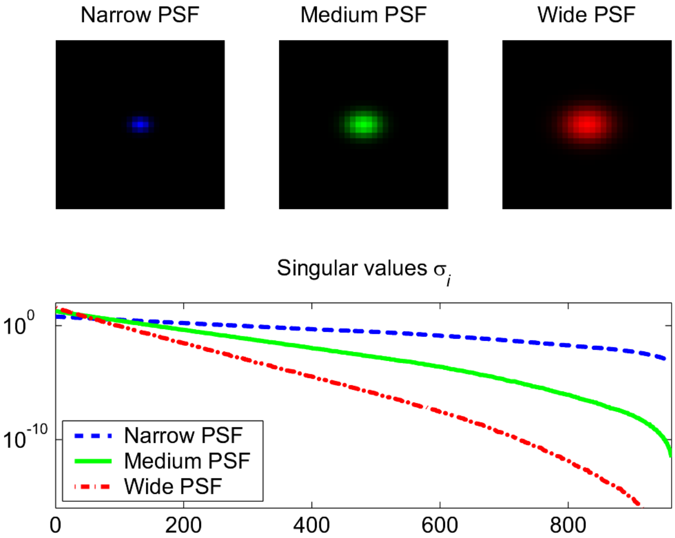
Recall that spectral filtering amounts to computing

$$\mathbf{x}_{\text{filt}} = \sum_{i=1}^{N} \phi_i \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i .$$

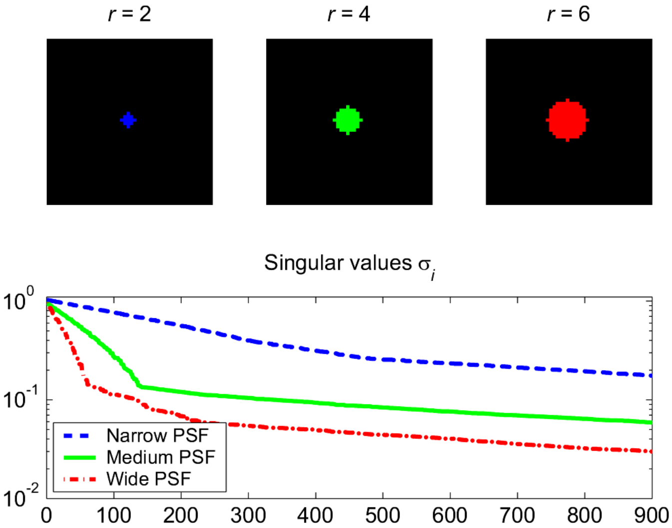
Hence we want to understand the behavior of the ingredients in this expression:

- How do the singular values σ_i depend on the PSF?
- How do the SVD coefficients $\mathbf{u}_i^T \mathbf{b}$ behave?
- What do the spectral basis vectors $\mathbf{v}_i$ look like?

The Singular Values: Gaussian PSF



The Singular Values: Out-of-Focus Blur

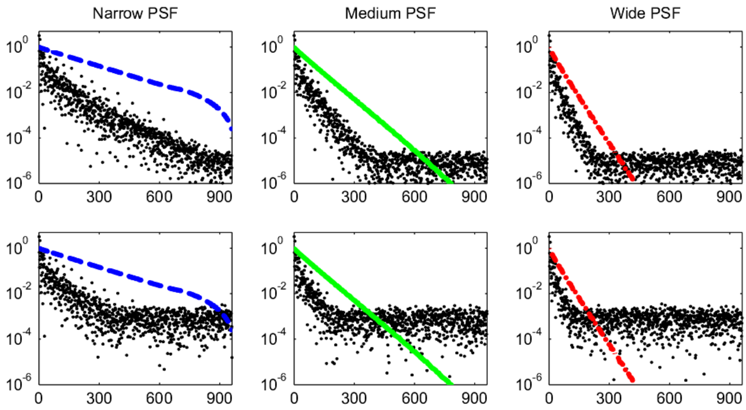


The Singular Values: Facts

1. As the blurring gets worse – i.e., the PSF gets “wider” – the singular values decay faster.
2. Even for narrow PSFs with a slow decay in singular values, the condition number $\text{cond}(\mathbf{A}) = \sigma_1/\sigma_N$ becomes large for large images.
3. The decay also depends on the “smoothness” of the PSF – the smoother, the faster the decay.
4. At one extreme, when the PSF consists of a single nonzero pixel, the matrix $\mathbf{A}$ is the identity and all singular values are identical (and the condition number of the matrix is one).
5. In the other extreme, when the PSF is so wide that $\mathbf{A}$ becomes the constant image, all but one of the singular values are zero.

The SVD Coefficients = Black Dots

The Gaussian PSF again: behavior of the coefficients $\mathbf{u}_i^T \mathbf{b}$.



Noise levels: $\|\mathbf{E}\|_\mathrm{F} = 3 \cdot 10^{-3}$ (top) and $\|\mathbf{E}\|_\mathrm{F} = 3 \cdot 10^{-1}$ (bottom).

The SVD Coefficients: Facts

1. Initially, the coefficients $|\mathbf{u}_i^T \mathbf{b}|$ decay – at a rate that is slightly faster than that of the singular values.
2. Later the coefficients level off at a plateau determined by the level of the noise in the image.
3. Coefficient that are larger in absolute value than the noise level carry information about the data.
4. Coefficients at the noise level are dominated by the noise, hiding the true information.
5. In other words,

$$\mathbf{u}_i^T \mathbf{b} \approx \begin{cases} \mathbf{u}_i^T \mathbf{b}_{\mathrm{exact}} & \text{for small } i \\ \mathbf{u}_i^T \mathbf{e} & \text{for large } i. \end{cases}$$

When Spectral Filtering Works

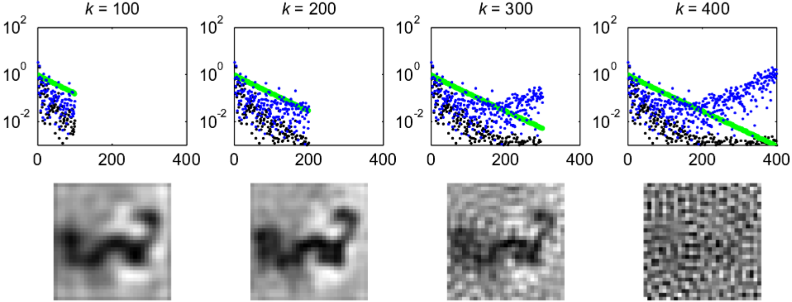
For any spectral filtering method

$$\mathbf{x}_{\text{filt}} = \sum_{i=1}^{N} \phi_i \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i .$$

we must choose the filters ϕ_i so that the information in the initial coefficients dominates the filtered solution.

- The index where the transition between the two types of behavior in $\mathbf{u}_i^T \mathbf{b}$ occurs, depends on the noise level and the decay of the unperturbed coefficients.
- Over-smoothing: If we include too few terms, then we miss available information in the data.
- Under-smoothing: If we include too many terms, then we include noisy components.

Illustration of TSVD



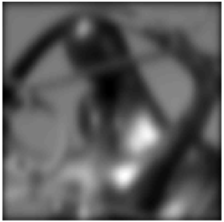
Top: singular values σ_i (green solid curve), right-hand side coef- ficients $|\mathbf{u}_i^T \mathbf{b}|$ (black dots) and TSVD solution coefficients $|\mathbf{u}_i^T \mathbf{b}/\sigma_i|$ (blue dots) for $k = 200$, 300 and 400 using the medium blur.

Bottom: the corresponding TSVD reconstructions.

The Decline and Fall . . .

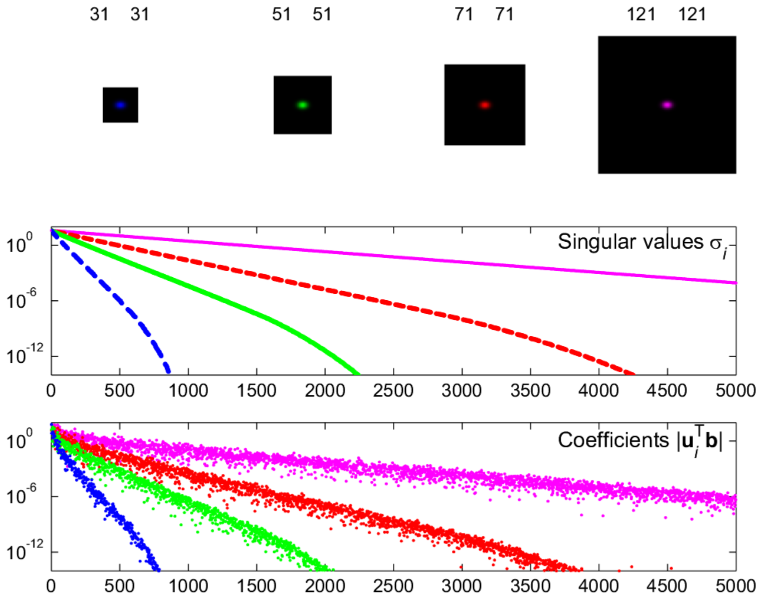
We illustrate the change in the decay of the singular values σ_i and the coefficients $\mathbf{u}_i^T\mathbf{b}$ as the size of the image increases.

The exact test image $\mathbf{X}_{\text{exact}}$ and the blurred image $\mathbf{B}$:



We also use four smaller sub-images from the central part of $\mathbf{X}_{\text{exact}}$.

The four images are blurred with a Gaussian PSF with the same parameters s_1 and s_2.


31 31
51 51
71 71
121 121
Singular values σ_i
10^0
10^-6
10^-12
0 500 1000 1500 2000 2500 3000 3500 4000 4500 5000
Coefficients |u_i^T b|
10^0
10^-6
10^-12
0 500 1000 1500 2000 2500 3000 3500 4000 4500 5000

The Importance of the Discrete Picard Condition

- The decay of the singular values σ_i becomes slower as the image size increases, for a fixed PSF.
- For all four test problems the coefficients $\mathbf{u}_i^T \mathbf{b}$ decay – on average – faster than the corresponding singular values.
- This behavior is an intrinsic property of inverse problems, known as the *discrete Picard condition*.
- Hence we also see – on average – a decay (perhaps a slight decay only) of the absolute values of the SVD coefficients $\mathbf{u}_i^T \mathbf{b}/\sigma_i$ for the naive solution.
- Consequently, it is the initial SVD coefficients that are dominated by the exact data.

Very Important Points from SVD Analysis

1. All singular values decay gradually to zero, and the typical behavior is that the wider the PSF and the larger the size of the image, the slower the decay.
2. The SVD coefficients $|\mathbf{u}_i^T \mathbf{b}|$ satisfy the discrete Picard condition, i.e., they decay (on average) faster than the singular values.
3. The spectral components which are large in absolute value primarily contain pure data, while those with smaller absolute value are dominated by noise.
4. The former components typically correspond to the larger singular values.

One remaining thing to study: the basis vectors $\mathbf{v}_i$ ...

Basis Vectors and Basis Images

We go back and forth between an image array and its vector representation via the "vec" notation:

$$\mathbf{x}_{\text{filt}} = \text{vec}(\mathbf{X}_{\text{filt}}) \qquad \text{and} \qquad \mathbf{v}_i = \text{vec}\big(\mathbf{V}^{[i]}\big), \quad i = 1, \ldots, N.$$

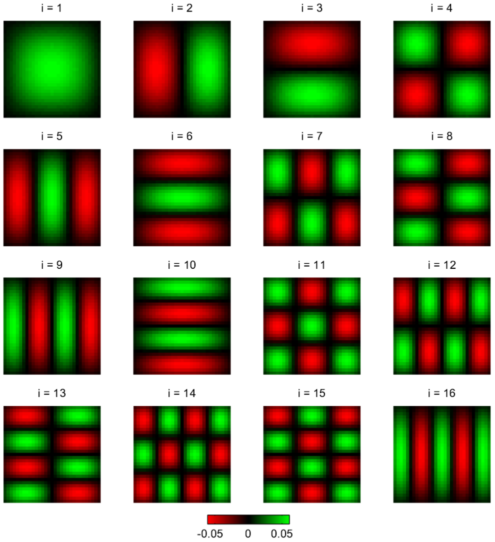
$\mathbf{X}_{\text{filt}}$ is the 2D representation of the filtered solution $\mathbf{x}_{\text{filt}}$, and the matrices $\mathbf{V}^{[i]}$ are the 2D representations of the singular vectors $\mathbf{v}_i$.

Using these quantities, we can write the filtered solution image as

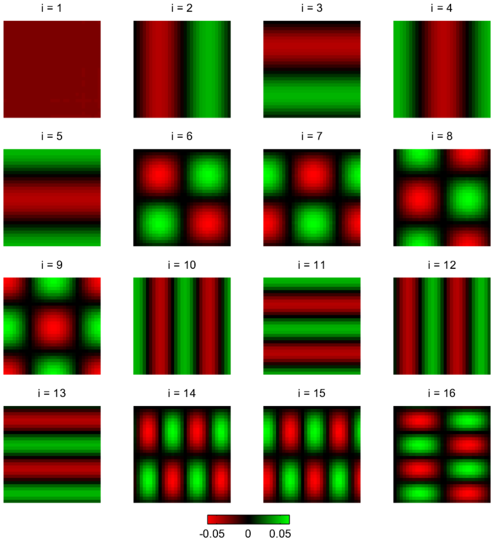
$$\mathbf{X}_{\text{filt}} = \sum_{i=1}^{N} \phi_i \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{V}^{[i]}.$$

So what do the basis matrices $\mathbf{V}^{[i]}$ look like?

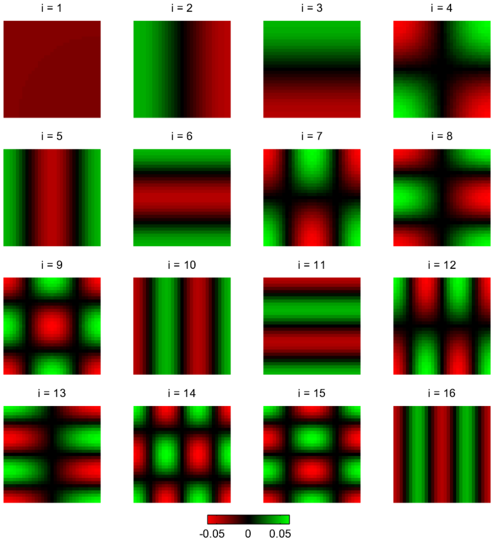
SVD Basis Images, Gaussian PSF, Zero BC



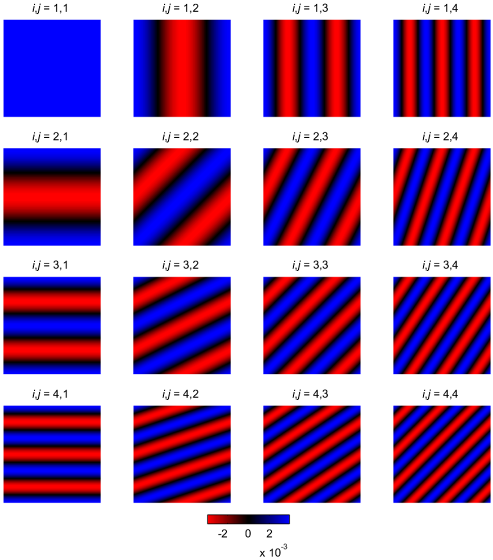
SVD Basis Images, Gaussian PSF, Periodic BC



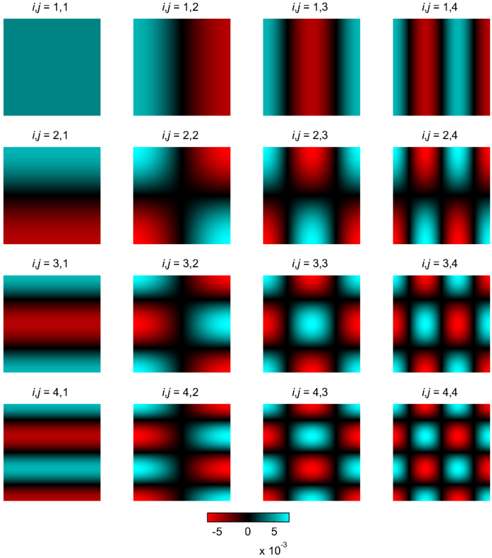
SVD Basis Images, Gaussian PSF, Reflexive BC



FFT Basis Images (Independent of the PSF)



DCT Basis Images (Independent of the PSF)



Summary of Basis Image Properties

1. The basis images (or spectral basis components) $\mathbf{V}^{[i]}$ become more oscillatory as the index i increases.
2. The basis images satisfy the boundary conditions.
3. The SVD basis images depend on the PSF, while the FFT and DCT basis images are independent of the PSF.
4. Each FFT basis image is characterized by one spatial frequency and one "angle."
5. Each DCT basis image is characterized by two spatial frequencies.