STA5103 Selected Topics in Frontiers of Statistics

Homework 4

12112627 李乐平

Question 1

Answer:

Now we want to solve robust PCA problem using ADMM:

$$\min_{X\in\mathbb{R}^{m\times n}} \|X\|_* + \rho\|M-X\|_1$$

Formalize the robust PCA problem into ADMM-suitable form:

$$\min_{X\in\mathbb{R}^{m\times n}} \|X\|_* + \rho\|Z\|_1,\ \text{s.t. } X+Z=M$$

Step 1: Augmented Lagrangian

$$L(X,Z,Y) = \|X\|_* + \rho\|Z\|_1 + \frac{\mu}{2}\left\|X+Z-M+\frac{Y}{\mu}\right\|_F^2$$

where $Y\in\mathbb{R}^{m\times n}$ is the dual variable, $\mu>0$ is the penalty parameter, and $\|\cdot\|_F$ is the Frobenius norm.

Step 2: Update X

$$X_{k+1} = \arg\min_X \|X\|_* + \frac{\mu}{2}\left\|X+Z_k-M+\frac{Y_k}{\mu}\right\|_F^2$$

This is a singular value thresholding (SVT) problem. The solution is:

$$X_{k+1} = D_{\frac{1}{\mu}}(M-Z_k-\frac{Y_k}{\mu})$$

where $D_\tau(A)$ is the singular value thresholding operator defined as:

$$D_\tau(A) = US_\tau(\Sigma)V^T, \text{ with } A=U\Sigma V^T \text{ (SVD of } A) \text{ and } S_\tau(\sigma_{ii}) = \max(\sigma_{ii}-\tau,0)$$

Step 3: Update Z

$$Z_{k+1} = \arg\min_Z \rho\|Z\|_1 + \frac{\mu}{2}\left\|X_{k+1}+Z-M+\frac{Y_k}{\mu}\right\|_F^2$$

This is a soft-thresholding problem for entry-wise L_1-norm. The solution is:

$$Z_{k+1} = S_{\frac{\rho}{\mu}}(M-X_{k+1}-\frac{Y_k}{\mu}), S_\tau(a) = \text{sgn}(a)\cdot\max(|a|-\tau,0)$$

$S_\tau(a)$ applies soft-thresholding to each entry of the matrix.

Step 4: Update Y

$$Y_{k+1} = Y_k + \mu(X_{k+1}+Z_{k+1}-M)$$

Step 5: Iterate until convergence

Repeat the above steps until convergence, typically when:

$$\|X_{k+1}+Z_{k+1}-M\|_F \le \varepsilon$$

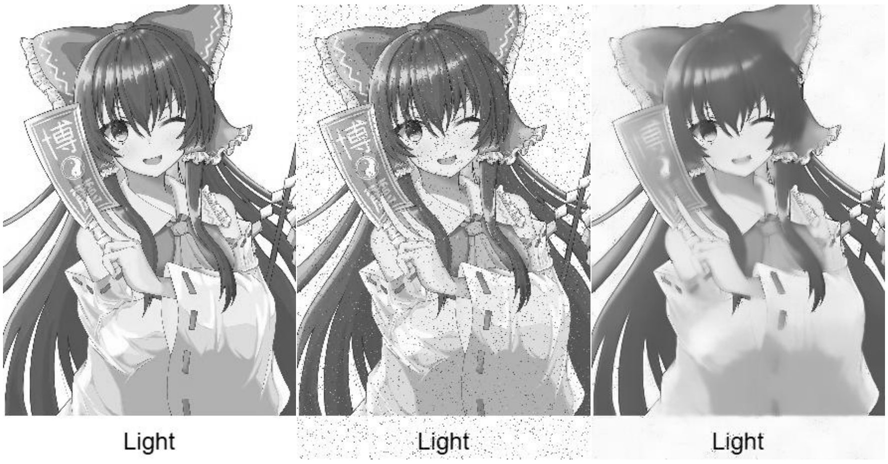
where ε is a small tolerance.

Robust PCA has various applications in image processing, particularly in tasks where the data is corrupted by outliers or noise. For instance, in image denoising, robust PCA can separate the low-rank background (representing clean images) from sparse outliers (representing noise). It is also used in background subtraction in video processing, where the goal is to separate the moving foreground objects (sparse) from the static background (low-rank). Additionally, in image compression, robust PCA can be applied to decompose an image into its low-rank approximation, helping to compress the image by retaining only the significant components, while discarding sparse or noise-like details. These methods are highly effective in handling imperfect or incomplete data, making them widely used in real-world image restoration and analysis tasks.

Question 2

Answer:

Please refer to `nl_mean.py` to see my implementation of nonlocal mean.



Original Image Noisy Image Denoised Image

Question 3

Answer:

Deep Image Prior (DIP) is a unique framework in image processing that utilizes the structure of convolutional neural networks (CNNs) as a prior for solving various image reconstruction tasks. Unlike conventional approaches, DIP does not rely on pre-trained models or large datasets. Instead, it demonstrates that the architectural biases of an untrained CNN can capture the statistical properties of natural images. By directly training the network on a degraded image, DIP exploits this innate structure to reconstruct high-quality images through a process of iterative optimization.

The architecture of DIP typically involves an encoder-decoder structure with skip connections. This design is significant because it facilitates the representation of complex image features while maintaining the spatial coherence of the original image. The encoder extracts hierarchical features, and the decoder reconstructs the image from these features, often guided by the skip connections that preserve low-level details. The input to the network is generally a random tensor, such as Gaussian noise, which is gradually transformed into the target output image through optimization.

The training process in DIP is distinctive for its simplicity and focus. The network parameters are optimized to minimize a loss function that compares the network’ s output with the degraded image, such as pixel-wise differences for tasks like denoising, inpainting, or super-resolution. This process resembles traditional overfitting; however, it leverages the CNN's architectural bias, which prioritizes generating clean, structured outputs rather than noise. Early stopping is crucial in this framework, as prolonged training can lead the network to model noise present in the corrupted input. This characteristic ensures that the output image reflects the underlying structure of the original image before significant noise overfitting occurs.

Recent advancements have extended the capabilities of DIP, integrating techniques like neural tangent kernel (NTK) analysis to deepen our understanding of its optimization dynamics ("Analysis of Deep Image Prior and Exploiting Self-Guidance for Image Reconstruction"). Researchers have also explored sparse and compact architectures to improve efficiency and reduce overfitting risks ("Chasing Better Deep Image Priors between Over- and Under-parameterization"). These enhancements illustrate how DIP continues to evolve, addressing its limitations while expanding its applications in medical imaging, compressive sensing, and other domains ("Deep Random Projector: Accelerated Deep Image Prior").