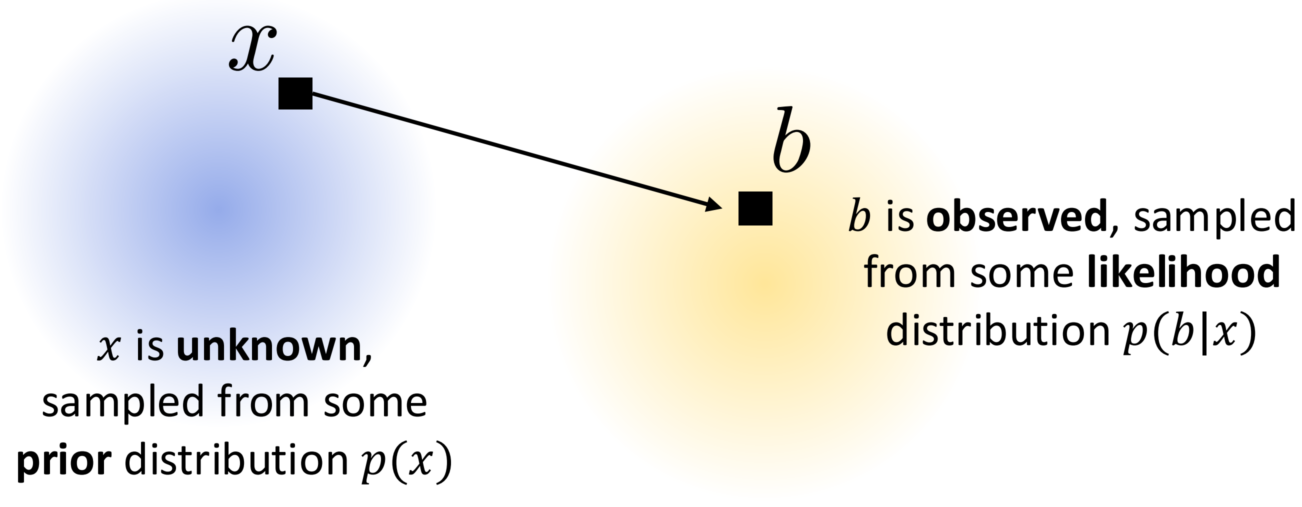
Problem Session 6

Topics

- Bayesian perspective on inverse problems
- 3 Optimization Methods
  - Gradient Descent with Adam
  - Half-Quadratic Splitting
  - ADMM
- Implementation Tips

A Bayesian Perspective on Inverse Problems



We want to recover x given b

$$x_{\text{MAP}} = \operatorname{argmin}_x \left(-\log p(b|x) - \log p(x)\right)$$

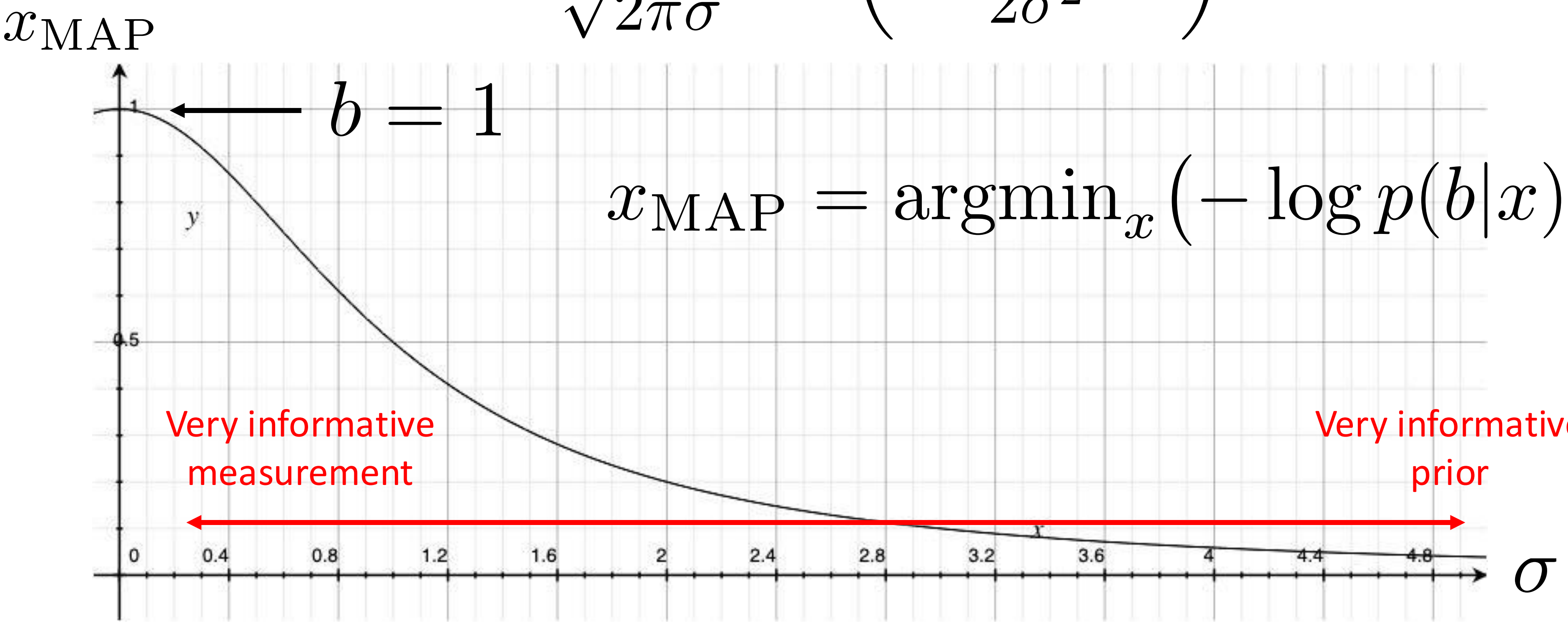
1D Example

e.g.

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right)$$ 0-mean Gaussian prior

$$p(b|x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(b-x)^2}{2\sigma^2}\right)$$ Gaussian Noise (variance σ^2)

$$x_{\mathrm{MAP}} = \mathrm{argmin}_x \left(-\log p(b|x) - \log p(x)\right)$$



Linear Inverse Problems

$$x_{\mathrm{MAP}} = \mathrm{argmin}_x \left(-\log p(b|x) - \log p(x) \right)$$

e.g.

$$-\log p(b|x) = \frac{1}{2\sigma^2} \|Ax - b\|_2^2$$ Linear model with Gaussian noise

$$-\log p(x) = \lambda \|Dx\|_1$$ TV regularization

Linear Inverse Problems

$$\text{minimize} \ \frac{1}{2\sigma^2}\|Ax - b\|_2^2 + \lambda\|Dx\|_1 \quad \text{over } x$$

3 ways to tackle an optimization problem

- Look for a closed-form expression of the minimizer (e.g. with FOC) → Fast but does not always work and requires an explicit prior
- Use a gradient-based optimization method (e.g. Adam) → Simple but can be slow, does not always converge and requires an explicit prior (Task 1)
- Use something smarter (e.g. HQS, ADMM) → Can be faster, more robust and quite versatile… Topic of tasks 2 and 3!

Task 1 – Gradient Descent

$$\text{minimize } \frac{1}{2}\|Cx-b\|_2^2+\lambda\text{TV}(x) \quad \text{over } x$$

better: isotropic

$\sqrt{(" _0 !)_1^2 + (" _3 !)_1^2}$

easier: anisotropic

$\sqrt{(" _0 !)_1^2} + \sqrt{(" _3 !)_1^2}$

!

Task 1 – Gradient Descent

- Your task is to
  - ❑ Implement the computation of the TV norm (hint: compute gradients in Fourier space)
  - ❑ Run GD with ADAM and report qualitative + quantitative results
  - ❑ Compare isotropic and anisotropic TV

Task 2 - HQS

$$\text{minimize } f(x) + g(Dx) \quad \text{over } x$$

↓ Formulate as a constrained optimization problem

$$\text{minimize } f(x) + g(z)$$

$$\text{subject to } Dx = z$$

↓ `while not converged`

$$x \leftarrow \text{argmin}_x \left(f(x) + \frac{\rho}{2} \|Dx - z\|_2^2 \right)$$ "x-update"

$$z \leftarrow \text{argmin}_z \left(g(z) + \frac{\rho}{2} \|Dx - z\|_2^2 \right)$$ "z-update"

We have 2 optimization problems instead of 1!

We can find closed-form solution to these…

Task 2 - HQS

$$\text{minimize } f(x) + g(Dx) \quad \text{over } x$$

Deconvolution with TV

$$f(x) = \|Cx - b\|_2^2$$

$$g(Dx) = \lambda \|Dx\|_1$$

D is the gradient operator

Task 2 - HQS

x-update

$$f(x) = \|Cx - b\|_2^2 \qquad\qquad g(Dx) = \lambda \|Dx\|_1$$

$$x \leftarrow \mathrm{argmin}_x \left(f(x) + \frac{\rho}{2} \|Dx - z\|_2^2 \right) = \mathrm{prox}(z)$$

$$\mathrm{prox}_{\|\cdot\|_{!},\mathrm{B}}(\boldsymbol{z}) = \mathcal{F}^{9:} \left\{ \frac{\mathcal{F}\{c\}^* \cdot \mathcal{F}\{b\} + \rho\left(\mathcal{F}\{d_0\}^* \cdot \mathcal{F}\{z_{:}\} + \mathcal{F}\{d_3\}^* \cdot \mathcal{F}\{z_2\}\right)}{\mathcal{F}\{c\}^* \cdot \mathcal{F}\{c\} + \rho\left(\mathcal{F}\{d_0\}^* \cdot \mathcal{F}\{d_0\} + \mathcal{F}\{d_3\}^* \cdot \mathcal{F}\{d_3\}\right)} \right\}$$

See lecture notes for derivation

Task 2 - HQS

z-update

$$f(x) = \|Cx - b\|_2^2 \qquad g(Dx) = \lambda \|Dx\|_1$$

$$z \leftarrow \text{argmin}_z \left(g(z) + \frac{\rho}{2} \|Dx - z\|_2^2 \right) = \text{prox}(v = Dx)$$

$$\text{prox}_{\|\cdot\|_1,\rho}(\boldsymbol{v}) = \mathcal{S}_\kappa(\boldsymbol{v}) = \begin{cases} v - \kappa & v > \kappa \\ 0 & |v| \leq \kappa \\ v + \kappa & v < -\kappa \end{cases} = (v - \kappa)_+ - (-v - \kappa)_+$$

$$\kappa = \lambda/\rho$$

The z-update is already implemented ☺

Task 2 - HQS

Deconvolution with DnCNN

$$f(x) = \|Cx - b\|_2^2$$

$$g(Dx) = \lambda\psi(x)$$

D is the identity matrix here

What is ψ?

ψ is the prior implicitly represented by a Gaussian denoiser

Task 2 - HQS

ψ is the prior implicitly represented by a Gaussian denoiser

We have a denoiser $\mathcal{D}$, which we assume behaves as a MAP estimator

i.e., we assume there exists ψ such that

$$\mathcal{D}\left(x, \sigma^2\right) = \text{argmin}_z \left(\frac{1}{2\sigma^2}\|z - x\|^2 + \psi(z)\right)$$

Task 2 - HQS

x-update

$$f(x) = \|Cx - b\|_2^2 \qquad g(Dx) = \lambda\psi(x)$$

$$x \leftarrow \text{argmin}_x \big(f(x) + \frac{\rho}{2}\|x - z\|_2^2\big) = \text{prox}(z)$$

$$\text{prox}_{\|\cdot\|_2,\rho}(\boldsymbol{z}) = \mathcal{F}^{-1}\left\{\frac{\mathcal{F}\{c\}^* \cdot \mathcal{F}\{b\} + \rho\mathcal{F}\{z\}}{\mathcal{F}\{c\}^* \cdot \mathcal{F}\{c\} + \rho}\right\}$$

Task 2 - HQS

z-update

$$f(x) = \|Cx - b\|_2^2 \qquad g(Dx) = \lambda\psi(x)$$

$$z \leftarrow \operatorname{argmin}_z \left(\psi(z) + \frac{\rho}{2\lambda}\|x - z\|_2^2\right) = \operatorname{prox}(x)$$

$$\operatorname{prox}(x) = \mathcal{D}\left(x, \sigma^2 = \frac{\lambda}{\rho}\right)$$

The z-update is already implemented ☺

Task 2 - HQS

Takeaways

- HQS transforms an inverse problem with **no closed-form solution** into **two optimization problems**.
- Most often, these subproblems have **closed form solutions**.
- A denoiser can be seen as a closed form solution to a MAP problem. Therefore, **a denoiser implicitly defines a prior**.

Task 3 – Under-Constrained Problem

Single-pixel imaging

$$b = Ax$$

A is a “flat” matrix (more columns than rows).

This is an under-constrained problem.

Idea #1: Make the problem well-posed with “least-norm”.

Task 3 – Least-Norm

Idea #1: Make the problem well-posed with “least-norm”.

$$\text{minimize } \|x\|_2^2$$
$$\text{subject to } Ax = b$$

We have a closed-form solution:

$$x = A^T(AA^T)^{-1}b$$

Task 3 – Least-Norm

$$x = A^T(AA^T)^{-1}b$$

- Compute $u = (AA^T)^{-1}b$
- Compute $x = A^T u$

Use `cg` and `LinearOperator` from `scipy.sparse.linalg`

Task 3 – Least-Norm

```
LinearOperator(size, matvec=method)
```

1. Size = (number of measurements, number of measurements)
2. Method = a function that implements the operation AA^T

(You are given Afun and Atfun already)

```
cg(A=LinearOperator, b=b, maxiter=num_iters, tol=cg_tolerance)
```

Don't forget to apply the final A^T!

Task 3 – ADMM

Single-pixel imaging

$$b = Ax$$

Idea #2: Add a prior.

$$\text{minimize } \frac{1}{2}\|Ax - b\|_2^2 + \lambda\psi(Dx)$$

There is no closed form solutions anymore ☹

Task 3 – ADMM

$$\text{minimize } f(x) + g(Dx) \quad \text{over } x$$

↓ Formulate as a constrained optimization problem

$$\text{minimize } f(x) + g(z)$$

$$\text{subject to } Dx = z$$

↓ `while not converged`

$$x \leftarrow \text{argmin}_x \Big(f(x) + \frac{\rho}{2} \| Dx - z + u \|_2^2 \Big)$$ "x-update"

$$z \leftarrow \text{argmin}_z \Big(g(z) + \frac{\rho}{2} \| Dx - z + u \|_2^2 \Big)$$ "z-update"

$$u \leftarrow u + Dx - z$$ "dual-update"

Task 3 – ADMM

$$\underset{x}{\text{minimize}} \quad \frac{1}{2}\|\mathbf{Ax}-\mathbf{b}\|^2 + \lambda \text{TV}(\mathbf{x})$$

$$\mathbf{x} \leftarrow \underset{x}{\text{argmin}}\, \frac{1}{2}\|\mathbf{Ax}-\mathbf{b}\|^2 + \frac{\rho}{2}\left\|\mathbf{Dx}-\mathbf{z}+\mathbf{u}\right\|^2$$

$$\mathbf{z} \leftarrow \underset{\mathbf{z}}{\text{argmin}}\, \lambda\|\mathbf{z}\|_p + \frac{\rho}{2}\left\|\mathbf{Dx}-\mathbf{z}+\mathbf{u}\right\|^2$$

$$\mathbf{u} \leftarrow \mathbf{u} + \mathbf{Dx} - \mathbf{z}$$

- Use Conjugate Gradients for the x-update. Combine the least squares objectives first.

Task 3 – ADMM

- Sum of squared losses is also a squared loss:

$$\|Ax - b\|^2 + \lambda\|Cx - d\|^2 = \left\| \begin{bmatrix} A \\ \sqrt{\lambda}C \end{bmatrix} x - \begin{bmatrix} b \\ \sqrt{\lambda}d \end{bmatrix} \right\|^2$$

- Least squares analytic solution

$$\text{minimize} \quad \|Ax - b\|^2 \quad \longrightarrow \quad x = (A^T A)^{-1} A^T b$$

Note: this is the least-square problem, the “over-constrained” version of “Ax=b”

Task 3 – ADMM

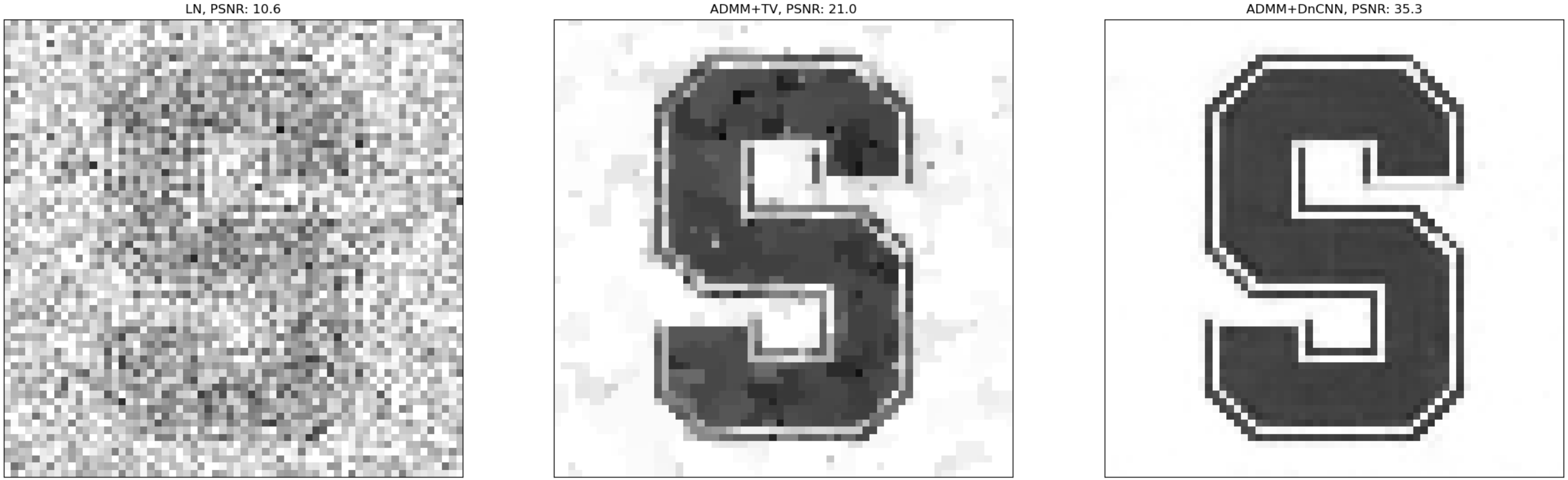
$$\underset{x}{\text{minimize}} \quad \frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2 + \lambda\Psi_{\mathrm{DnCNN}}(\mathbf{x})$$

$$\mathbf{x} \leftarrow \underset{x}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2 + \frac{\rho}{2}\left\|\mathbf{x}-\mathbf{z}+\mathbf{u}\right\|^2$$

$$\mathbf{z} \leftarrow \mathrm{DnCNN}(\mathbf{x}+\mathbf{u})$$

$$\mathbf{u} \leftarrow \mathbf{u}+\mathbf{x}-\mathbf{z}$$

- Use Conjugate Gradients to do the x-update. Combine the least squares objectives first.


LN, PSNR: 10.6
ADMM+TV, PSNR: 21.0
ADMM+DnCNN, PSNR: 35.3

Summary

- Inverse problems:
  - Deconvolution with a prior
  - Single-pixel imaging (under-constrained with a prior)
- Priors:
  - TV (anisotropic, isotropic)
  - Denoisers

Good luck with the homework!