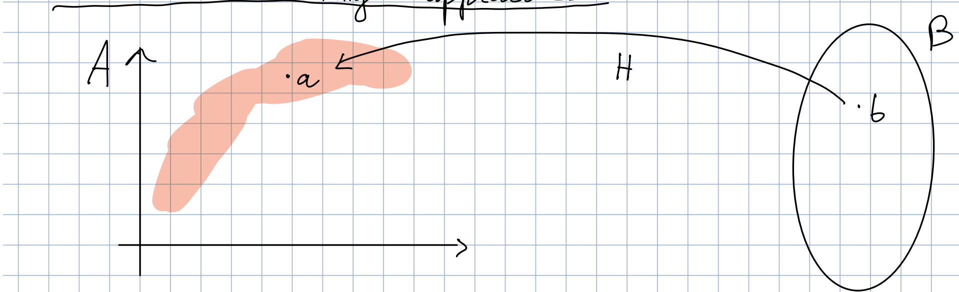
3. Machine learning approaches

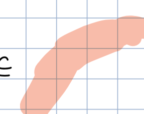


H: depends on
- F
- The prior on a (R)

H: can be (partially) learned

Choices to make:

i) <u>Training set</u>

supervised: $\{(b_i, a_i)\}_{i=1}^N$, $b_i = F(a_i) + \varepsilon_i$ | Unsupervised $\{a_i\}_{i=1}^N \subseteq$  $[\{b_i\}_{i=1}^N]$

ii) Architecture / model for $H = H_\theta$

<u>End-to-end</u> (supervised)

Ignore F, learn H directly by minimizing

$$\bigstar \qquad \min_\theta \frac{1}{N} \sum_{i=1}^N \| H_\theta(b_i) - a_i \|^2 + \mathrm{Reg}(\theta)$$

H_θ: NN, θ are parameters

→ nothing but a regression problem in sup. learning

<u>Plug-and-play</u> (unsupervised: $\{a_i\}$)

Recall:

$$a_0 = 0$$

prox op.

$$a_{j+1} = S\left(a_j - \gamma F^T(F a_j - b)\right)$$

stop: $a_K = H(b)$

Key obs: S is a denoiser:

$$S_f(a) = \operatorname{argmin}_x \frac{1}{2}\|x - a\|^2 + f(x)$$

Learn S! How? Minimize ★ with a training set:

$$\{(a_j + \tilde{\varepsilon}_j , a_j)\}_{j=1}^N$$

Unrolling / Unfolding (supervised)

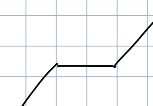
prox op.

$$a_0 = 0$$

$$a_{j+1} = S\left(a_j - \gamma F^T(F a_j - b)\right)$$

stop: $a_K = H(b)$

e.g. $S = \Phi^T S_\lambda \Phi$ S_λ



$H : b \longmapsto a_K$ is a NN with S_λ as nonlinearity

Key idea: replace S at iteration j by a NN N_{θ_j}

The full H_θ is:

is a NN (shallow) itself

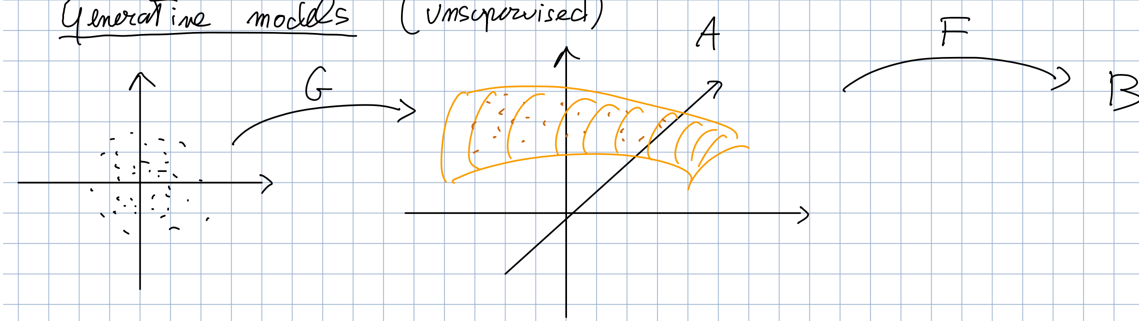
$$\begin{cases} a_0 = 0 \\ a_{j+1} = N_{\theta_j}\left(a_j - \gamma F^T(F a_j - b)\right) \\ H_\theta(b) = a_K \end{cases}$$

Train H_θ, $\theta = (\theta_0, \theta_1, \dots, \theta_{K-1})$, by minimizing ★

$$\min_\theta \frac{1}{N} \sum_{i=1}^N \|H_\theta(b_i) - a_i\|^2 + \mathrm{Reg}(\theta)$$

The training depends on F!

Generative models (unsupervised)



Instead of solving the IP:

$$F(a) = b \qquad\qquad a = G(z)$$

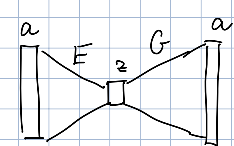
Solve

$$F(G(z)) = b$$

Then $\quad a_{rec} = G(z_{rec})$

Many options:

- variational autoencoders



$G \circ E \approx I$

- GANs
- diffusion models

Deep image priors (no learning, no training set)

Given a NN

$$G_\theta : \mathbb{R}^d \longrightarrow A = \mathbb{R}^n$$

(fix an architecture). Sample $z \sim N(0, I)$ in $\mathbb{R}^d$ (fix it!)

Solve

$$\min_\theta \; \| F(G_\theta(z)) - b \|_B^2 + \operatorname{Reg}(\theta)$$

$a = G_\theta(z)$, z fixed.

References

- Arridge - Maass - Öktem - Schönlieb, Acta Numerica 2019
- Ongie - Jalal - Metzler - Baraniuk - Dimakis - Willet, 2020
- Ulyanov - Vedaldi - Lempitsky, CVPR 2018
- Bora - Jalal - Price - Dimakis, PMLR 2017
- Adler - Öktem, IP 2017
- Bubba (ed.), De Gruyter, 2025