Solving Imaging Inverse Problems using GAN Priors

Viraj Shah

March 28, 2022

ECE Department, UIUC

Parts of this work have been done in collaboration with Dr. Chinmay Hegde, and presented at ICASSP 2018, ICASSP 2019.
An extended version is under review in an IEEE Journal.

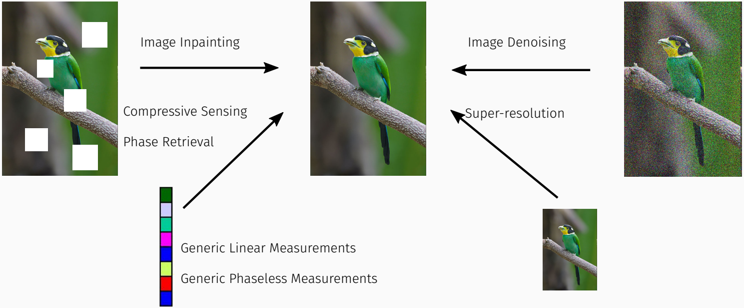
Introduction

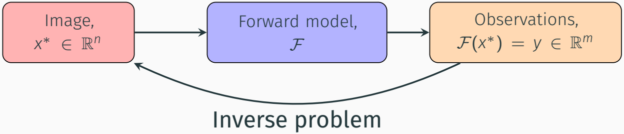
Forward Problem: *given quantity, determine observations*;
Inverse Problem: *given observations, recover quantity.*

Example
Quantity: mass of the earth;
Observations: earth's gravitational field.

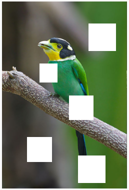
Imaging, optics, astrophysics, and seismic geo-exploration,...

Introduction: Inverse problems in imaging





Goal is to recover x^*, given the $\mathcal{F}$ and y.



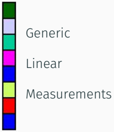


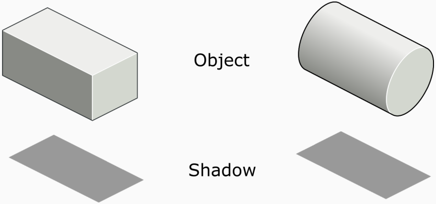
Image Denoising
$\mathcal{F}(x) = I_n x + e$

Image Inpainting
$\mathcal{F}(x) =$ pixel-wise selection

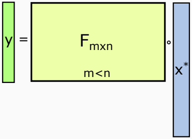
Compressive Sensing
$\mathcal{F}(x) = F_{m \times n} x,\ m << n$

Super-resolution
$\mathcal{F}(x) =$ lowpass + downsampling

Inverse Recovery is ill-posed



Compressive Sensing
(with Gaussian measurements)



$F \in \mathbb{R}^{m\times n}$; **with** $m < n$**, infinite many solutions.**

Additional information about x^* is required!

Assume: x^* exhibits a *structure* that is known apriori

Solve a constrained optimization problem:

$$\begin{aligned} \widehat{x} = \operatorname*{arg\,min} \; & \|y - Fx\|_2^2, \\ \text{s. t.} \quad & x \in \mathcal{S}. \end{aligned} \tag{1}$$

$\mathcal{S}$ is a set of all images that obeys the given prior

Related Work

Related Work: Various priors

Hand-crafted Priors

Total variation [ROF92; Cha04]

Sparsity [Can+06]

Structured sparsity [HIS15]

Dictionaries [EA06]

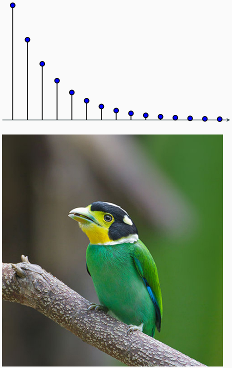
Learned Priors

End-to-end learning [MPB15]

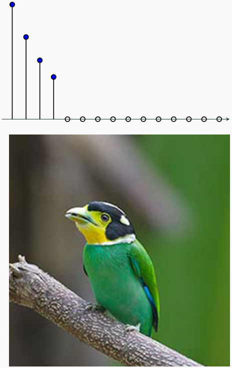
Generative priors [Bor+17]

Deep Image Priors [UVL18]

Related Work: Hand-crafted prior - Sparsity



Original image



Sparse image (DCT)

Key Assumption

Manifold of natural images $\approx$ Set of $s-$sparse images, $\mathcal{S}$

Related Work: Sparse recovery

$\boxed{\text{The prior } \mathcal{S} = \{x \in \mathbb{R}^n \mid \|x\|_0 \leq s\}}$

Minimize l_0−norm:

$$\min_x \|x\|_0, \quad s.t. \quad Fx = y, \tag{2}$$

Perfect recovery if $m \geq O(s\log(n/s))$ [FR13]

Can be solved using Lasso [Tib96], basis pursuit [CDS01],...

$\boxed{\text{Shown success on variety of linear and non-linear inverse problems}}$

Denoising [Don95], **Super-resolution** [Don+11]
Image inpainting [XS10], **Phase retrieval** [JH17]

What if $m < O(s\log(n/s))$ **?**

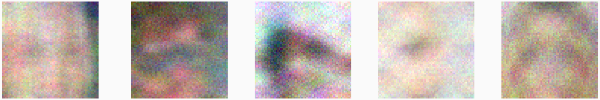
Related Work: Limitations of sparsity

Original $n = 12288$ | $m = 50$ | $m = 100$ | $m = 200$ | $m = 500$ | $m = 1000$ | $m = 5000$

Sparse reconstructions using Lasso (in DCT domain)

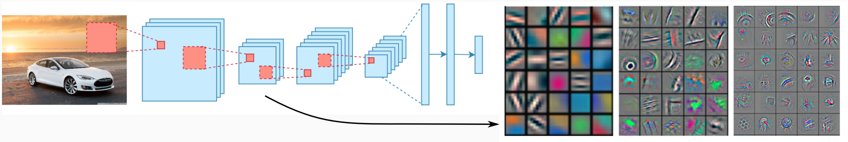
Poor performance for low values of m

Can we leverage similar images from existing datasets?



These are sparse in known basis, but do they resemble natural images?

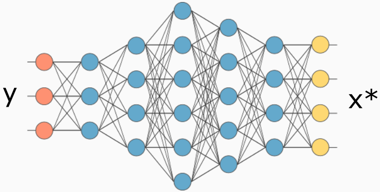
Poor discrimination!



CNNs are known to learn rich image representations.

Learn a mapping from y to x^*

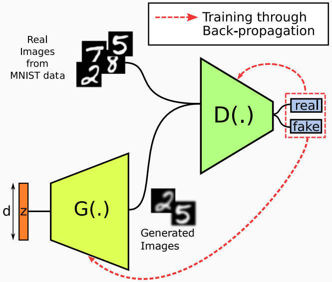
Several approaches for different inverse problems
[Kul+16; MPB15; MB17; Don+16; KKLML16]

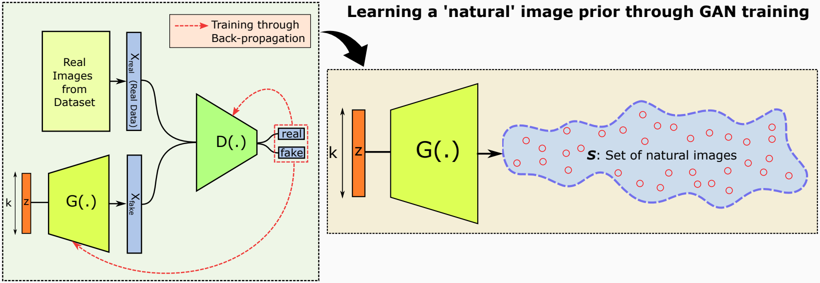


Each new problem requires fresh training!

How to learn natural image manifold directly?

Generative Adversarial Networks (GANs)
[Goo+14; RMC16]





New Assumption

Manifold of natural images ≈ Range of well-trained Generator (G), $\mathcal{S}$

The prior $\mathcal{S} = \{x \in \mathbb{R}^n \mid x = G(z),\ z \in \mathbb{R}^k\}$

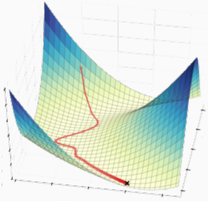
CSGM Algorithm [Bor+17]:

- Obtain a well-trained Generator G.
- Replace $x = G(z)$ in Eq. 1, and solve:

$$\hat{z} = \operatorname*{arg\,min}_{z \in \mathbb{R}^k} \|y - FG(z)\|_2^2, \quad \hat{x} = G(\hat{z}) \tag{3}$$

Recovery is possible if $m \geq O(kd \log n)$ for appropriate G

How to solve? G is non-convex!



Use Gradient Descent with Backprop

- Search remains limited to $\mathcal{S}$
- May get stuck in local minima
- Requires multiple random restarts
- No convergence guarantees

Contributions

We propose Projected Gradient Descent to **explore the space outside $\mathcal{S}$**

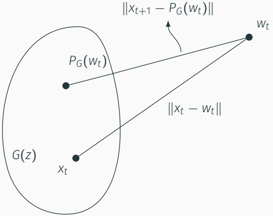
Step 1: Unconstrained Exploration

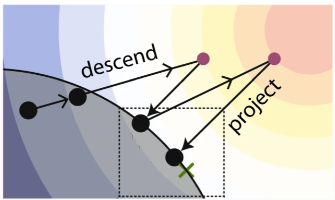
GD update on $\|(y - Fx)\|_2^2$:

$$w_t \leftarrow x_t + \eta_{gd} F^T(y - Fx_t)$$

Step 2: Projection

$$x_{t+1} = P_G(w_t) := G\left(\arg\min_z \|w_t - G(z)\|_2^2\right),$$





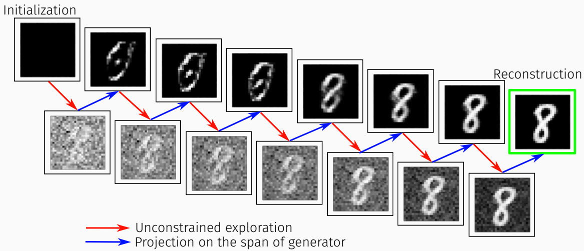
How to project on range of G ?

[1] Image from [Bol+17]

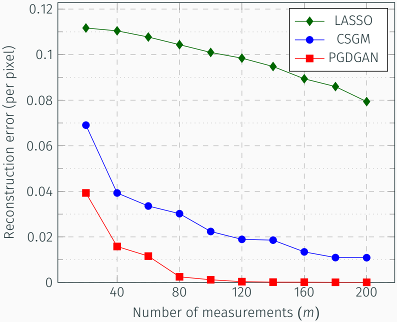
CONTRIBUTIONS: PGDGAN ALGORITHM [SH18]

How to project on range of G ?

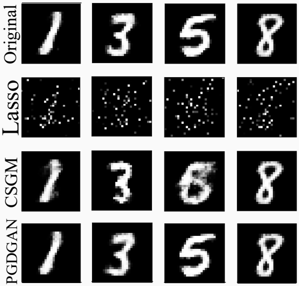
Multiple options: simplest is to use GD via backprop.



- Empirical results on MNIST, celebA
- Improved performance
- Random restarts are not needed
- Provable linear convergence
- Extended to non-linear problems as well
- Extension to handle model mismatch



(a)



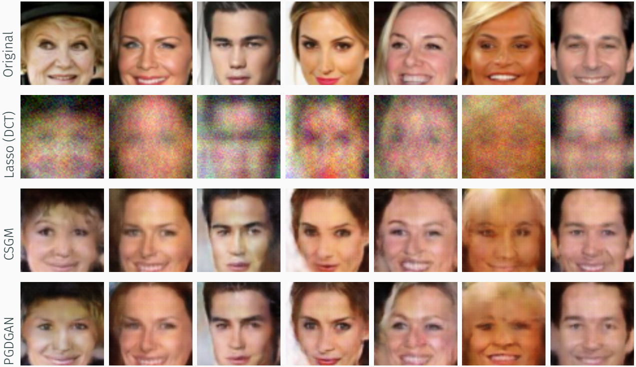
(b)

Comparisons among PGDGAN [SH18], CSGM [Bor+17], and Lasso-DCT [Tib96]

Reconstruction results on MNIST [LeC+98]:

- Generator is fully-connected network with $k = 20$.
- Image dimension is $n = 784$
- Reconstruction with $m = 100$ measurements (in (b)).

Contributions: Experimental results on celebA



Reconstruction results on CelebA [Liu+15]:

- DCGAN with $k = 100$.
- Unseen test images.
- Reconstruction with $m = 1000$ measurements.

Theorem (Guarantee: linear convergence)

Under certain conditions on F and m, the sequence (x_t) *defined by the PGDGAN algorithm with converges to* x^* *with high probability.* $\psi(\cdot)$ *is* l_2-*norm.*

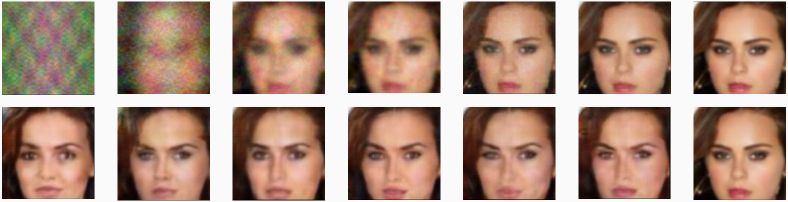
$$\psi(x_{t+1}) \leq \left(\frac{1}{\eta_{gd}\gamma} - 1\right)\psi(x_t)$$

Key ideas:

- The difference of any two images in $\mathcal{S}$ lies away from nullspace of F. $(S - REC)$.
- Spectral norm of F is upper-bounded by $\sqrt{\gamma}$.
- $P_G(\cdot)$ is a orthogonal projection operator.
- The learning rate obeys: $\frac{1}{2\gamma} < \eta_{gd} < \frac{1}{\gamma}$

Challenges and Recent Developments

What if target image is not in the range of G? $\implies$ Representation Error



Sparse reconstructions using Lasso and PGDGAN

$\boxed{\text{Performance of GAN prior saturates due to limited capacity of Generator}}$

(a) Glow model [KD18],

(b) Progressive-GAN model [Kar+17].

$\boxed{\text{Use better GAN variants}}$

Use both GAN prior and Sparsity

$$\boxed{x^* = x_b^* + x_i^*;\ x_b^* \text{ comes from G, } x_i^* \text{ is sparse}}$$

$$\min_{z, x_i} \|x_i\|_1 \quad \text{s.t.} \quad F(G(z) + x_i) = y,$$

Solve using alternating minimization [DGE18; Sha19]

Also useful in case of model mismatch

$\neq$

GAN Trained

on Dog Images

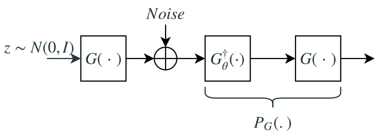
Inference on

Bird Image

Target image belongs to completely different image distribution!

How to obtain perfect projection on range of G ?

Learn the projection operator



[RLB19]

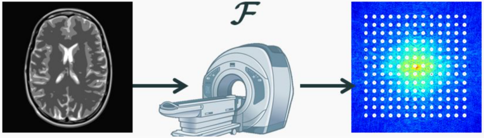
Invert G layer by layer

- Exact recovery is possible under certain conditions [Lei+19]

Challenge 3: Complex-valued representations

Many inverse problems operate in Complex domain

MRI, Phase Unwrapping, Speech etc.



Requires Complex-valued GANs

Deep Learning in Complex domain is challenging

- Design complex-valued network components [Tra+18; Vas+20]
  - activation function, softmax, batchnorm
- Incorporate Fourier Domain
- May lead to superior performance

Summary and Future Work

Summary:

- Replacing hand-crafted priors with learned priors
- Novel PGD-based algorithms with theoretical guarantees
- Works for variety of linear and non-linear inverse problems

Future directions:

- Complex-valued neural networks
- More challenging inverse problems such as modulo imaging

References

E. Bolduc et al. "Projected gradient descent algorithms for quantum state tomography". In: *npj Quantum Inf* (2017).

A. Bora et al. "Compressed Sensing using Generative Models". In: *Proc. Int. Conf. Machine Learning* (2017).

E. Candès et al. "Compressive sampling". In: *Proc. of the intl. congress of math.* Vol. 3. Madrid, Spain. 2006, pp. 1433–1452.

S. Chen, D. Donoho, and M. Saunders. "Atomic decomposition by basis pursuit". In: *SIAM review* 43.1 (2001), pp. 129–159.

A. Chambolle. "An algorithm for total variation minimization and applications". In: *Jour. of Math. imaging and vision* 20.1 (2004), pp. 89–97.

M. Dhar, A. Grover, and S. Ermon. "Modeling Sparse Deviations for Compressed Sensing using Generative Models". In: *Proc. Int. Conf. Machine Learning*. 2018.

References II

W. Dong et al. “Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization”. In: *IEEE Trans. Image Processing* 20.7 (2011), pp. 1838–1857.

C. Dong et al. “Image super-resolution using deep convolutional networks”. In: *IEEE Trans. Pattern Anal. Machine Intell.* 38.2 (2016), pp. 295–307.

D. Donoho. “De-noising by soft-thresholding”. In: *IEEE Trans. Inform. Theory* 41.3 (1995), pp. 613–627.

M. Elad and M. Aharon. “Image denoising via sparse and redundant representations over learned dictionaries”. In: *IEEE Trans. Image Processing* 15.12 (2006), pp. 3736–3745.

S. Foucart and H. Rauhut. *A mathematical introduction to compressive sensing*. Vol. 1. 3. Springer, 2013.

I. Goodfellow et al. “Generative adversarial nets”. In: *Proc. Adv. in Neural Processing Systems (NIPS)*. 2014, pp. 2672–2680.

C. Hegde, P. Indyk, and L. Schmidt. “Fast Algorithms for Structured Sparsity”. In: *Bulletin of the EATCS* 1.117 (2015), pp. 197–228.

G. Jagatap and C. Hegde. “Fast Sample-Efficient Algorithms for Structured Phase Retrieval”. In: *Proc. Adv. in Neural Processing Systems (NIPS)*. 2017.

Tero Karras et al. “Progressive growing of gans for improved quality, stability, and variation”. In: *arXiv preprint arXiv:1710.10196* (2017).

Durk P Kingma and Prafulla Dhariwal. “Glow: Generative flow with invertible 1x1 convolutions”. In: *Advances in Neural Information Processing Systems*. 2018, pp. 10215–10224.

J. Kim, J. Kwon Lee, and K. Mu Lee. “Accurate image super-resolution using very deep convolutional networks”. In: *Proc. IEEE Conf. Comp. Vision and Pattern Recog. (CVPR)*. 2016, pp. 1646–1654.

K. Kulkarni et al. “Reconnet: Non-iterative reconstruction of images from compressively sensed measurements”. In: *Proc. IEEE Conf. Comp. Vision and Pattern Recog. (CVPR)*. 2016, pp. 449–458.

Y. LeCun et al. “Gradient-based learning applied to document recognition”. In: *Proc. of the IEEE* 86.11 (1998), pp. 2278–2324.

Q. Lei et al. “Inverting Deep Generative models, One layer at a time”. In: *Proc. Adv. in Neural Processing Systems (NIPS)*. 2019.

Z. Liu et al. “Deep learning face attributes in the wild”. In: *Proc. of the IEEE Intl. Conf. on Comp. Vision*. 2015, pp. 3730–3738.

A. Mousavi and R. Baraniuk. “Learning to invert: Signal recovery via deep convolutional networks”. In: *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)* (2017).

A. Mousavi, A. Patel, and R. Baraniuk. “A deep learning approach to structured signal recovery”. In: *Proc. Allerton Conf. Communication, Control, and Computing*. 2015, pp. 1336–1343.

A. Raj, Y. Li, and Y. Bresler. “GAN-based Projector for Faster Recovery in Compressed Sensing with Convergence Guarantees”. In: *arXiv preprint arXiv:1902.09698* (2019).

A. Radford, L. Metz, and S. Chintala. “Unsupervised representation learning with deep convolutional generative adversarial networks”. In: *Proc. Int. Conf. Learning Representations (ICLR)* (2016).

References V

- L. Rudin, S. Osher, and E. Fatemi. “Nonlinear total variation based noise removal algorithms”. In: *Physica D: Nonlinear Phenomena* 60.1-4 (1992), pp. 259–268.
- V. Shah and C. Hegde. “Solving Linear Inverse Problems using GAN Priors: An Algorithm with Provable Guarantees”. In: *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*. 2018.
- V. Shah. “Algorithms for Solving Inverse Problems Using Generative Models”. In: *MS Dissertation, Iowa State University* (2019).
- Robert Tibshirani. “Regression shrinkage and selection via the lasso”. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1 (1996), pp. 267–288.
- C. Trabelsi et al. “Deep Complex Networks”. In: *ArXiv:1705.09792* (2018).
- Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. “Deep image prior”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 9446–9454.

B. Vasudeva et al. “Co-VeGAN: Complex-Valued Generative Adversarial Network for Compressive Sensing MR Image Reconstruction”. In: *ArXiv:2002.10523* (2020).

Z. Xu and J. Sun. “Image inpainting by patch propagation using patch sparsity”. In: *IEEE Trans. Image Processing* 19.5 (2010), pp. 1153–1165.

Appendix

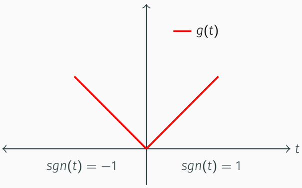
Experimental results for Super-resolution

Original images are in top row. Middle row depicts downsampled images (4×), and bottom row shows reconstruction.

Extension to phase retrieval problem

Phase retrieval problem:

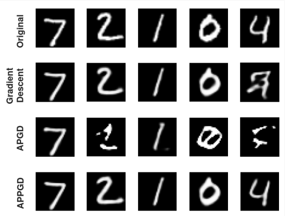
$$\widehat{x} = \arg\min_{x} \|y - |Fx|\|^2$$
$$\text{s.t.} \quad x = G(z),$$



$g(t) = \mathrm{abs}(t)$

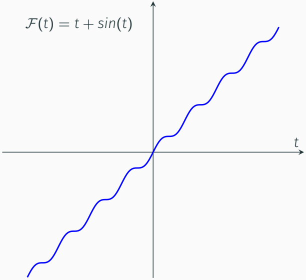
Alternating Phase PGD:

for t = 1,...T **do**
  $p_{t-1} \leftarrow sgn(Fx_{t-1})$
  $w_{t-1} \leftarrow x_{t-1} + \eta F^T(y \odot p_{t-1} - Fx_{t-1})$
  $x_t \leftarrow \mathcal{P}_G(w_{t-1}) = G\left(\arg\min_z \|w_{t-1} - G(z)\|\right)$
end for
$\widehat{x} \leftarrow x_T$



We provide empirical results for two non-linear inverse problems.

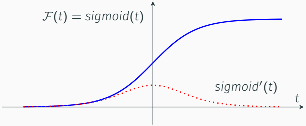
1. **Sinusoidal model**, with $\mathcal{F}(x) = Fx + sin(Fx)$.
   - We use l_2-loss as ψ.



2. **Sigmoid model**, with $(x) = sigmoid(Fx) = \frac{1}{1+\exp(-Fx)}$.
   - We use a loss function specified as:

$$\psi(t) = \frac{1}{m}\sum_{i=1}^{m}\left(\Theta(f_i^T t) - y_i f_i^T t\right),$$

where, $\Theta(\cdot)$ is integral of $\mathcal{F}(\cdot)$, and f_i represents the rows of the measurement matrix F.

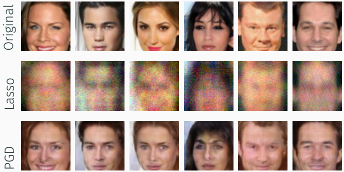


The gradient of the loss:

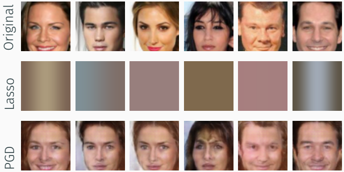
$$\nabla\psi(t) = \frac{1}{m}F^T(sigmoid(Ft) - y).$$

Experimental Results: Non-linear Inverse Problems

- **We perform the experiments on CelebA Dataset:**
  - We use DCGAN with both G and D are CNNs with 4 hidden layers each.
  - Dimensions of the input z is $k = 100$.
  - Test images are kept unseen during training.
  - Total number of updates is set to 1000, with $T = 10$ and $T_{in} = 100$.
  - We reconstruct the images with $m = 1000$ measurements.

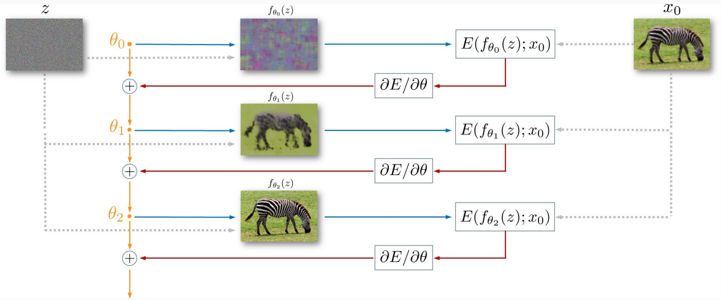


(a) Sinusoidal model;



(b) Sigmoid model.

Deep Image Prior: randomized neural network used as a handcrafted prior [UVL18]



Solve following through neural network training:

$$\theta^* = \arg\min_{\theta} \|f_\theta(z) - x_0\|_2^2; \quad x^* = f_\theta(z).$$

The difference vector of any two images in the set $\mathcal{S}$ should lie away from the nullspace of the matrix F.

S-REC (Set Restricted Eigenvalue Condition)

Let $\mathcal{S} \in \mathbb{R}^n$. F is $m \times n$ matrix. For parameters $\gamma > 0,\ \delta \geq 0$, matrix F is said to satisfy the S-REC$(\mathcal{S}, \gamma, \delta)$ if,

$$\|F(x_1 - x_2)\|^2 \geq \gamma \|x_1 - x_2\|^2 - \delta,$$

for $\forall x_1, x_2 \in \mathcal{S}$.

Theorem (Guarantee: linear convergence)

Let G be a generator with range $\mathcal{S}$. F is satisfying the S-REC$(\mathcal{S}, \gamma, \delta)$ with probability $1 - p$, and has $\|Fv\| \leq \rho \|v\|$ for every $v \in \mathbb{R}^n$ with probability $1 - q$. $\rho^2 \leq \gamma$.
Then, for every $x^* \in \mathcal{S}$, the sequence (x_t) defined by the PGDGAN algorithm converges to x^* with probability at least $1 - p - q$.

We extend the above algorithm for non-linear inverse problems:

- We generalize the loss function to be $\psi(\cdot)$ and the projection oracle to P_G.
- Assume that the ψ has a continuous gradient $\nabla\psi = \left(\frac{\partial\psi}{\partial x_i}\right)_{i=1}^{n}$.
- We define the ε-approximate projection oracle P_G as,

Approximate projection

A function $P_G : \mathbb{R}^n \to \mathrm{Range}(G)$ is an ε-approximate projection oracle if for all $x \in \mathbb{R}^n$, $P_G(x)$ obeys:

$$\|x - P_G(x)\|_2^2 \leq \min_{z\in\mathbb{R}^k} \|x - G(z)\|_2^2 + \varepsilon.$$

$\epsilon-$PGD Algorithm:

- Initialization: $x_0 \leftarrow \mathbf{0}$
- Gradient update step: $w_t \leftarrow x_t - \eta\nabla\psi(x_t)$
- Projection step: $x_{t+1} \leftarrow P_G(w_t)$

$\epsilon-$ PGD: Theoretical Results

Theorem (Linear Convergence of $\epsilon-$PGD)

Under certain conditions on ψ, ϵ-PGD algorithm convergences linearly up to a ball of radius $O(\gamma\Delta) \approx O(\varepsilon)$.

$$\psi(x_{t+1}) - \psi(x^*) \leq \left(\frac{\beta}{\alpha} - 1\right)(\psi(x_t) - \psi(x^*)) + O(\varepsilon)\,.$$

The analysis for linear problem is a special case of the above theorem.

Proved using:

- ψ follows Restricted Strong Convexity/Smoothness conditions with constants α, β.
- Gradient at the minimizer is small:$\|\nabla\psi(x^*)\|_2 \leq \gamma$
- Range of G is compact: $\mathrm{diam}(\mathrm{Range}(G)) = \Delta$.
- $\gamma\Delta \leq O(\varepsilon)$.
- $1 \leq \frac{\beta}{\alpha} < 2$

$\epsilon-$ PGD: Theoretical Results

We introduce more general restriction conditions on the $\psi(\cdot)$:

Restricted Strong Convexity/Smoothness

Assume that ψ satisfies $\forall x, y \in S$:

$$\frac{\alpha}{2}\|x-y\|_2^2 \leq \psi(y) - \psi(x) - \langle \nabla\psi(x), y-x \rangle \leq \frac{\beta}{2}\|x-y\|_2^2.$$

for positive constants α, β.

Theorem (Linear Convergence of $\epsilon-$PGD)

If ψ satisfies RSC/RSS over Range(G) with constants α and β, then ε-PGD algorithm convergences linearly up to a ball of radius $O(\gamma\boldsymbol{\Delta}) \approx O(\varepsilon)$.

$$\psi(x_{t+1}) - \psi(x^*) \leq \left(\frac{\beta}{\alpha} - 1\right)(\psi(x_t) - \psi(x^*)) + O(\varepsilon)\,.$$

The analysis for linear problem is a special case of the above theorem.