Practical work on image restoration

Convex regularisation and contour aware image restoration

Prerequisite: prepare in advance the answer to the questions no. 1 through 5.

The present practical work is devoted to image deconvolution. Let us consider a scene captured through a real observation system (of limited resolution that is): the image of a point is not a point but a spot. The observed image is blurred because it results from the superposition of the spots generated by each point from the real scene. Knowing the shape of the spot, *i.e.,* the point spread function of the imaging system, and the blurred image, the practical work seeks to recover the underlying sharp image, that is to invert the blurring. A previous work has been devoted to the Wiener-Hunt method and the present one proposes an extension in order to improve image resolution.

1 Wiener-Hunt Solution

We use the following mathematical model to describe the acquisition process:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{e}$$

where the vector $\boldsymbol{y}$ represents the observed data (the blurred image), the vector $\boldsymbol{x}$ represents the unknown true image (the sharp image), $\boldsymbol{H}$ is the convolution matrix and $\boldsymbol{e}$ is the vector accounting for the measurement and modelling errors.

To regularise the deconvolution problem, we include additional information regarding the spatial regularity of the true image. The most straightforward approach consists in introducing a penalty of the difference between the grey level of neighbouring pixels. We define the following penalized criterion:

$$\mathcal{J}_{\mathrm{Q}}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 + \mu \sum_{p\sim q} \varphi_{\mathrm{Q}}(x_p - x_q) \tag{1}$$

with $\varphi_{\mathrm{Q}}(\delta) = \delta^2$. We can rewrite the above criterion as follows:

$$\mathcal{J}_{\mathrm{Q}}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 + \mu\, \|\boldsymbol{D}\boldsymbol{x}\|^2 \ ,$$

where $\boldsymbol{D}$ is a difference matrix. The reconstructed image $\widehat{\boldsymbol{x}}_{\mathrm{Q}}$ is taken as the minimizer of this criterion:

$$\widehat{\boldsymbol{x}}_{\mathrm{Q}} = \arg\min_{\boldsymbol{x}} \mathcal{J}_{\mathrm{Q}}(\boldsymbol{x})\,.$$

As we are dealing with a quadratic criterion in $\boldsymbol{x}$, we have an analytic expression for the minimizer:

$$\widehat{\boldsymbol{x}}_{\mathrm{Q}} = (\boldsymbol{H}^{\mathrm{t}}\boldsymbol{H} + \mu\, \boldsymbol{D}^{\mathrm{t}}\boldsymbol{D})^{-1}\boldsymbol{H}^{\mathrm{t}}\boldsymbol{y}\,. \tag{2}$$

By introducing a circulant approximation, we can diagonalise all the matrices involved in (2) in the Fourier basis. In turn, this enables us to compute the solution for a marginal computational cost, that is Wiener-Hunt solution. We shall further exploit this circulant approximation in the non-quadratic case dealt within this work.

Given a well chosen value for μ, we are able to truly see a deconvolution effect. That said, the resolution and the ability to restore sharp edges, to allow for abrupt changes in grey levels of the restored image, is limited. The object of this practical work is to overcome this limitation.

2 Convex Regularisation

2.1 Huber Potential Function

As a mean to further enhance the resolution and have better edge preserving properties, we reconsider the potential function φ, taking for example:

$$\varphi_{\mathrm{H}}(\delta) = \begin{cases} \delta^2 & \text{if } |\,\delta\,| \leq T \\ 2T\,|\,\delta\,| - T^2 & \text{if } |\,\delta\,| \geq T \end{cases}$$

which is referred to as the Huber potential function. It has a quadratic behaviour up to a given threshold T and a linear behaviour afterwards (see the following figure).

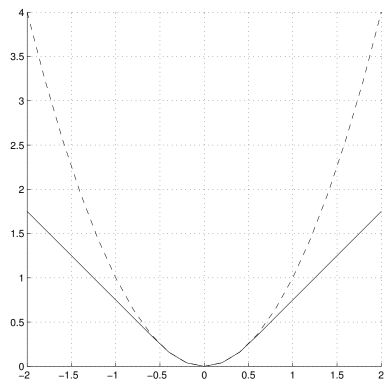

Figure 1: Huber function (threshold $T = 0.5$), solid line. Quadratic function, dotted line.

1. *Determine (theoritically) and trace (on your paper) the first two derivatives of φ_{H}. Trace as well the first and second derivative of φ_{Q} for comparison purposes.*

Starting from the Huber potential function φ_{H}, we define a new criterion similar to the one given in (1):

$$\mathcal{J}_{\mathrm{H}}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 + \mu \sum_{p \sim q} \varphi_{\mathrm{H}}(x_p - x_q)\,, \tag{3}$$

and we define the restored image $\widehat{\boldsymbol{x}}_{\mathrm{H}}$ as the minimizer of the above new criterion

$$\widehat{\boldsymbol{x}}_{\mathrm{H}} = \arg\min_{\boldsymbol{x}} \mathcal{J}_{\mathrm{H}}(\boldsymbol{x})\,. \tag{4}$$

2. *Intuition and qualitative analysis*

   2a. *Explain in your own words how this new potential function φ_{H} enables better edge preserving properties compared to the previous potential function φ_{Q}.*

   2b. *What is the influence of the threshold T?*

Remark 1 — *The chosen φ_{H} function is convex and so is the adequation to the data term. We deduce that the criterion $\mathcal{J}_{\mathrm{H}}$ as a whole is also convex: it thus has a unique minimum. We can also prove that it is strictly convex and that it has a unique minimizer.*

For an algorithmic reason that will be made explicit later on, we introduce an extra-parameter denoted α, a strictly positive real number. The penalty term is multiplied and divided by α so that:

$$\mathcal{J}_{\mathrm{H}}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 + \mu' \sum_{p \sim q} \alpha\, \varphi_{\mathrm{H}}(x_p - x_q)\,,$$

with $\mu' = \mu/\alpha$.

2.2 Optimisation

The problem we are facing now is how to actually compute the minimizer (4) of the new criterion (3). Unlike $\mathcal{J}_{\mathrm{Q}}$ given by (1), it is not quadratic: we do not have an analytic formula for its minimizer, as in (2). Nonetheless, it has a unique minimizer and there are numerical iterative algorithms allowing us to compute it, see for example [1–3]. In the following we shall study one of the available options.

2.2.1 Extended criterion and auxiliary variables

In order to reuse the previous result from the quadratic case under the circulant approximation, *i.e.,* the Wiener-Hunt solution, we introduce a new set of variables called *auxiliary variables*. More specifically, we introduce a variable a_{pq} for each pair of neighbouring pixels $(\,p\,,\,q\,)$ and we collect all these new variables in the vector $\boldsymbol{a}$.

We proceed with the construction of a so called *extended* criterion

$$\widetilde{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x},\boldsymbol{a}) = \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|^2 + \mu' \left[\sum_{p\sim q} \frac{1}{2}\left[(x_p - x_q) - a_{pq}\right]^2 + \widetilde{\zeta}_\alpha(a_{pq})\right] \tag{5}$$

which is a function of the unknown image $\boldsymbol{x}$ and of the auxiliary variables $\boldsymbol{a}$. It comprises three terms:

1. the least squares term (adequation to the data),
2. a quadratic term involving the differences between pixels, linked to the auxiliary variables, and
3. a term dealing only with the auxiliary variables involving a new function $\widetilde{\zeta}_\alpha$, called *auxiliary function*.

The keystone of the construction is to minimize the initial criterion (3) by minimizing the extended criterion (5) instead, that is to say, a desired property is:

$$\min_{\boldsymbol{a}} \widetilde{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x},\boldsymbol{a}) = \mathcal{J}_{\mathrm{H}}(\boldsymbol{x})\,. \tag{6}$$

In order to meet this property, reading (5), the function $\widetilde{\zeta}_\alpha$ can be clearly chosen so that:

$$\alpha\varphi(\delta) = \inf_a \left[\frac{1}{2}(\delta - a)^2 + \widetilde{\zeta}_\alpha(a)\right]\,. \tag{7}$$

Unsurprisingly, the design of $\widetilde{\zeta}_\alpha$ in order to satisfy this property is paramount. It relies on theoretical considerations as for example *convex duality* and *Legendre-Fenchel transform* (see [4–7], the lecture notes and the exercices) for more details. It is proved that $\widetilde{\zeta}_\alpha$ is a Huber function as well:

$$\widetilde{\zeta}_\alpha(a) = \alpha \begin{cases} \dfrac{1}{1-2\alpha}\,a^2 & \text{if } |\,a\,| \leq (1-2\alpha)\,T \\ 2T\,|\,a\,| - (1-2\alpha)\,T^2 & \text{if } |\,a\,| \geq (1-2\alpha)\,T \end{cases} \tag{8}$$

Remark 2 — *The parameter* $\alpha \in]0, 1/2[$ *enables a fine-tuning of the algorithm. It does not alter the solution itself, i.e., the restored image, instead it influences the inner workings of the optimisation algorithm, particularly its convergence speed.*

2.2.2 Algorithm

The results from the previous section, in particular equation (6), show that we can obtain the minimizer of $\mathcal{J}_{\mathrm{H}}(\boldsymbol{x})$ with respect to $\boldsymbol{x}$ by minimizing $\widetilde{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x},\boldsymbol{a})$ with respect to both $\boldsymbol{x}$ and $\boldsymbol{a}$:

$$\widehat{\boldsymbol{x}}_{\mathrm{H}} = \arg\min_{\boldsymbol{x}} \mathcal{J}_{\mathrm{H}}(\boldsymbol{x}) = \arg\min_{\boldsymbol{x}} \left\{\min_{\boldsymbol{a}} \widetilde{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x},\boldsymbol{a})\right\}\,.$$

From a practical stand point, we will compute the joint minimizer of $\widetilde{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{a})$ with respect to both $\boldsymbol{x}$ and $\boldsymbol{a}$:

$$(\widetilde{\boldsymbol{x}}_{\mathrm{H}}, \widetilde{\boldsymbol{a}}_{\mathrm{H}}) = \arg\min_{\boldsymbol{x},\boldsymbol{a}} \widetilde{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{a}) \,,$$

and we clearly have: $\widehat{\boldsymbol{x}}_{\mathrm{H}} = \widetilde{\boldsymbol{x}}_{\mathrm{H}}$.

3. *Take time to properly digest the news and make sense of it.*

Then, the computation of the joint minimizer of $\widetilde{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{a})$ with respect to $(\boldsymbol{x}, \boldsymbol{a})$, will be achieved by iterating a two stage process until convergence.

① Minimize $\widetilde{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{a})$ with respect to $\boldsymbol{x}$ for fixed $\boldsymbol{a}$; this yields $\overline{\boldsymbol{x}}(\boldsymbol{a}) = \arg\min_{\boldsymbol{x}} \widetilde{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{a})$.

② Minimize $\widetilde{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{a})$ with respect to $\boldsymbol{a}$ for fixed $\boldsymbol{x}$; in turn this yields $\overline{\boldsymbol{a}}(\boldsymbol{x}) = \arg\min_{\boldsymbol{a}} \widetilde{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{a})$.

4. *Give the explicit solution to ①. Show how it can be computed in an efficient manner as a Wiener-Hunt solution by using a circulant approximation. What's happen if $\boldsymbol{a} = \boldsymbol{0}$?*

5. *Regarding step ②.*

5a. *Examine criterion (5) and explain why it is possible to update the set of a_{pq}'s in an independent and parallel manner.*

5b. *This point makes explicit the update of any of the auxiliary variables a_{pq} as a function of the inter-pixel difference $\delta_{pq} = x_p - x_q$. To options are available.*

1. *By minimization of (7) given (8) and you should find*

$$\overline{a}_{pq} = \begin{cases} \delta_{pq} - 2\alpha T & \text{if } \delta_{pq} \geq T \\ (1-2\alpha)\,\delta_{pq} & \text{if } |\,\delta_{pq}\,| \leq T \\ \delta_{pq} + 2\alpha T & \text{if } \delta_{pq} \leq -T \end{cases}$$

2. *As expected from the lecture notes:*

$$\overline{a}_{pq} = \delta_{pq} - \alpha \varphi'_{\mathrm{H}}(\delta_{pq})$$

You could check that the two are identical. For your implementation, you can take it as it is.

The advantage of this half-quadratic approach is that both steps ① and ② are explicit, whereas the direct minimization of $\mathcal{J}_{\mathrm{H}}(\boldsymbol{x})$ is not. The algorithm can also be given in the following form.

- Initialize $\boldsymbol{a}^{[0]} = \boldsymbol{0}$
- For $k = 1, 2, \dots$ repeat
  1. Update $\boldsymbol{x}$: $\boldsymbol{x}^{[k]} = \arg\min_{\boldsymbol{x}} \widetilde{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{a}^{[k-1]}) = \dots$
  2. Update $\boldsymbol{a}$: $\boldsymbol{a}^{[k]} = \arg\min_{\boldsymbol{a}} \widetilde{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x}^{[k]}, \boldsymbol{a}) = \dots$

2.2.3 Practical implementation

This section deals with the practical implementation in *Matlab* and the analysis of the results.

6a. *Implement the optimisation method from the previous section. Take time to properly structure and comment your code.*

6b. *Compare the results to the ones obtained using the Wiener-Hunt method. Comment upon the influence of the two parameters of the method (μ and T).*

6c. *Analyse the results in the frequency domain as well.*

6d. *Analyse and comment upon the speed of convergence of the algorithm as a function of α. By trial and error determine a good value for α.*

Please, keep in mind that α does not impact the restored image itself but it only influences the optimisation algorithm, particularly its convergence speed.

3 Further analysis (optional) : interpretation in terms of line variables

This section proposes a new interpretation of the previous solution (and not a new solution). It is an alternative interpretation of the criterion (3) and its minimizer (4). It makes use of the so called *line variables* that, in a way, reveal the discontinuities in the reconstructed image. To this end, we introduce a second *extended* criterion:

$$\bar{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{\ell}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 + \mu \left[\sum_{p \sim q} \ell_{pq} \left(x_p - x_q \right)^2 + \sum_{p \sim q} \bar{\zeta}(\ell_{pq}) \right] . \tag{9}$$

The line variables $\ell_{pq} \in [0, 1]$ are unobserved and they are introduced between neighbouring pixels to break or attenuate inter-pixel interactions. The idea of using line variables, binary valued at the origin, was first introduced in the 1980s by [8] (see also [9, 10]).

7. *Intuition and qualitative analysis, a second time.*

7a. *Explain in your own words how this new criterion allows for a contour aware image restoration. What value of ℓ_{pq} allows for a strong discontinuity between neighbouring pixels (p, q) ?*

7b. *What happens if all ℓ_{pq} are equal to 0 ? And all equal to 1 ?*

7c. *What is the role of the $\bar{\zeta}$ function? What would happen if it were absent (i.e., if $\bar{\zeta}$ is the null fiction)?*

As before, the function $\bar{\zeta}$ can be constructed by making use of the *convex duality* framework, see for example [4–7] (and your lecture notes). As in the previous section, we directly give the result:

$$\bar{\zeta}(\ell) = s^2 \left(1/\ell - 1 \right)$$

and we will show that this choice enable to connect the extended criterion (9) to the initial one (3).

8a. *We introduce a function of ℓ denoted $\overline{\psi}_\delta(\ell)$ having δ as parameter: $\overline{\psi}_\delta(\ell) = \ell\delta^2 + \bar{\zeta}(\ell)$. Prove that by minimizing $\overline{\psi}_\delta(\ell)$ with respect to ℓ we obtain $\varphi_{\mathrm{H}}(\delta)$:*

$$\varphi_{\mathrm{H}}(\delta) = \min_{\ell} \overline{\psi}_\delta(\ell) .$$

We denote the minimizer by $\overline{\ell} = \theta(\delta) = \arg\min_\ell \overline{\psi}_\delta(\ell)$. Study and graph the function θ.

8b. *Deduce that the criterion (3) is obtained from the criterion (9) by minimizing the latter with respect to the auxiliary variables :*

$$\mathcal{J}_{\mathrm{H}}(\boldsymbol{x}) = \min_{\boldsymbol{\ell}} \bar{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{\ell})\,. \tag{10}$$

It should now be clear that $\mathcal{J}_{\mathrm{H}}(\boldsymbol{x})$ can be minimized with respect to $\boldsymbol{x}$ by minimizing $\widetilde{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{\ell})$ with respect to both $\boldsymbol{x}$ and $\boldsymbol{\ell}$. Do you confirm ?

Thus, the solution constructed by minimizing (9) with respect to both $\boldsymbol{x}$ and $\boldsymbol{\ell}$ is:

$$(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{\ell}}) = \underset{(\boldsymbol{x},\boldsymbol{\ell})}{\arg\min}\, \bar{\mathcal{J}}_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{\ell})$$

and it consists of the reconstructed image $\widehat{\boldsymbol{x}}$ and an image of the lines $\widehat{\boldsymbol{\ell}}$. According to (10), the reconstructed image $\widehat{\boldsymbol{x}}$ is the same as the one obtained in the previous part: $\widehat{\boldsymbol{x}} = \widehat{\boldsymbol{x}}_{\mathrm{H}}$.

9. *Draw the lines corresponding to the solution $\widehat{\boldsymbol{x}}_{\mathrm{H}}$.*

References

[1] J. Nocedal et S. J. Wright, *Numerical Optimization*, Series in Operations Research. Springer Verlag, New York, 2008.

[2] D. P. Bertsekas, *Nonlinear programming*, Athena Scientific, Belmont, MA, USA, 2nd edition, 1999.

[3] S. Boyd, N. Parikh, E. Chu, B. Peleato et J. Eckstein, *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*, Now Publishers Inc, Hanover, MA, USA, 2011.

[4] D. Geman et G. Reynolds, ¡¡ Constrained restoration and the recovery of discontinuities ¿¿, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, n° 3, pp. 367–383, mars 1992.

[5] D. Geman et C. Yang, ¡¡ Nonlinear image recovery with half-quadratic regularization ¿¿, *IEEE Trans. Image Processing*, vol. 4, n° 7, pp. 932–946, juillet 1995.

[6] P. Charbonnier, L. Blanc-Féraud, G. Aubert et M. Barlaud, ¡¡ Deterministic edge-preserving regularization in computed imaging ¿¿, *IEEE Trans. Image Processing*, vol. 6, n° 2, pp. 298–311, février 1997.

[7] J. Idier, ¡¡ Convex half-quadratic criteria and interacting auxiliary variables for image restoration ¿¿, *IEEE Trans. Image Processing*, vol. 10, n° 7, pp. 1001–1009, juillet 2001.

[8] S. Geman et D. Geman, ¡¡ Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images ¿¿, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, n° 6, pp. 721–741, novembre 1984.

[9] A. Blake et A. Zisserman, *Visual reconstruction*, The MIT Press, Cambridge, MA, USA, 1987.

[10] J. L. Marroquin, S. K. Mitter et T. A. Poggio, ¡¡ Probabilistic solution of ill-posed problems in computational vision ¿¿, *J. Amer. Stat. Assoc.*, vol. 82, pp. 76–89, 1987.