9

Two direct regularization methods

In this chapter we present two direct regularization methods, namely the Backus–Gilbert method and the maximum entropy regularization. Although these approaches have been designed for linear problems they can be applied to nonlinear problems as well.

9.1 Backus–Gilbert method

In the framework of Tikhonov regularization, the generalized inverse is not explicitly computed and is merely an analysis tool. The goal of the so-called mollifier methods is the computation of an approximate generalized inverse, which can then be used to obtain an approximate solution. Mollifier methods have been introduced by Louis and Maass (1990) in a continuous setting, and applied for discrete problems by Rieder and Schuster (2000).

To describe mollifier methods, we consider a semi-discrete Fredholm integral equation of the first kind

$$y_i = \int_0^{z_{\max}} k_i(z)\, x(z)\, \mathrm{d}z, \quad i = 1, \ldots, m, \tag{9.1}$$

and introduce a smoothing operator $A_\mu : X \to X$ by the relation

$$(A_\mu x)(z_0) = \int_0^{z_{\max}} a_\mu(z_0, z)\, x(z)\, \mathrm{d}z. \tag{9.2}$$

The parameter-dependent function a_μ in (9.2) is called mollifier and it is chosen such that $A_\mu x \to x$ as $\mu \to 0$ for all $x \in X$. Next, we assume that a_μ can be expressed as

$$a_\mu(z_0, z) = \sum_{i=1}^{m} k_i(z)\, k_{\mu i}^{\dagger}(z_0), \tag{9.3}$$

where $k_{\mu i}^{\dagger}$ are referred to as the contribution functions. In the framework of mollifier methods we choose a mollifier $\bar{a}_\mu$ and compute the contribution functions $k_{\mu i}^{\dagger}$ as the solution of

the constrained minimization problem

$$\min_{k_{\mu i}^{\dagger}} \int_0^{z_{\max}} \left[\bar{a}_\mu\left(z_0, z\right) - a_\mu\left(z_0, z\right)\right]^2 \, \mathrm{d}z \tag{9.4}$$

$$\text{subject to} \int_0^{z_{\max}} a_\mu\left(z_0, z\right) \, \mathrm{d}z = 1,$$

with a_μ being given by (9.3). The normalization condition in (9.4) just means that for $x \equiv 1$, $A_\mu x \equiv 1$ (cf. (9.2)). Once the contribution functions are known, we use the representation (cf. (9.1), (9.2) and (9.3))

$$\left(A_\mu x\right)\left(z_0\right) = \sum_{i=1}^{m} \left[\int_0^{z_{\max}} k_i\left(z\right) x\left(z\right) \, \mathrm{d}z\right] k_{\mu i}^{\dagger}\left(z_0\right) = \sum_{i=1}^{m} k_{\mu i}^{\dagger}\left(z_0\right) y_i, \tag{9.5}$$

to compute the mollified solution of the linear equation (9.1) with noisy data y_i^δ as

$$x_\mu^\delta\left(z_0\right) = \sum_{i=1}^{m} k_{\mu i}^{\dagger}\left(z_0\right) y_i^\delta. \tag{9.6}$$

Thus, in the framework of mollifier methods, instead of solving (9.1), we choose the mollifier and solve (9.3) with respect to the contribution functions as in (9.4). Equation (9.3) is also ill-posed as soon as equation (9.1) is, but the calculation of the mollified solution, according to (9.4) and (9.6), is expected to be a stable process because there are no errors in the data.

The transpose vector $\mathbf{k}_\mu^{\dagger T} = [k_{\mu 1}^{\dagger}, \ldots, k_{\mu m}^{\dagger}]$ reproduces the row vector of the generalized inverse $\mathbf{K}_\mu^{\dagger}$ corresponding to the altitude height z_0, and $a_\mu\left(z_0, z\right)$ can be interpreted as a continuous version of the averaging kernel matrix $\mathbf{K}_\mu^{\dagger}\mathbf{K}$.

The function $a_\mu\left(z_0, z\right)$ determines the resolution of the mollifier method at z_0, and for $x_\mu^\delta\left(z_0\right)$ to be meaningful, $a_\mu\left(z_0, z\right)$ should peak around z_0. To make $a_\mu\left(z_0, z\right)$ as localized as possible about the point z_0, we have to choose the mollifiers as smooth regular functions approximating a Dirac distribution. In fact, the choice of mollifiers depends on the peculiarities of the solution, and frequently used choices are (Louis and Maass, 1990)

$$\bar{a}_\mu\left(z_0, z\right) = \begin{cases} c, & |z - z_0| \le \mu, \\ 0, & \text{otherwise}, \end{cases}$$

$$\bar{a}_\mu\left(z_0, z\right) = c \, \mathrm{sinc}\left(\mu\left(z - z_0\right)\right),$$

$$\bar{a}_\mu\left(z_0, z\right) = c \exp\left(-\frac{\left(z - z_0\right)^2}{2\mu^2}\right),$$

where the parameter μ controls the width of the δ-like functions and c is a normalization constant.

Another variant of mollifier methods is the Backus–Gilbert method, also known as the method of optimally localized averages (Backus and Gilbert, 1967, 1968, 1970). In this approach, the averaging kernel function $a_\mu\left(z_0, z\right)$ is controlled by specifying a positive

δ^{-1}-like function $d_\mu(z_0, z)$ and then solving the constrained minimization problem

$$\min_{k^\dagger_{\mu i}} \int_0^{z_{\max}} d_\mu(z_0, z)\, a_\mu(z_0, z)^2 \, \mathrm{d}z \tag{9.7}$$
$$\text{subject to} \int_0^{z_{\max}} a_\mu(z_0, z) \, \mathrm{d}z = 1.$$

The function d_μ can be chosen as

$$d_\mu(z_0, z) = \left| \frac{z - z_0}{l} \right|^\mu \tag{9.8}$$

or as

$$d_\mu(z_0, z) = 1 - \exp\left(-\frac{1}{2} \left| \frac{z - z_0}{l} \right|^\mu \right), \tag{9.9}$$

where l is the correlation length and as before, μ is a parameter which controls the width of the δ^{-1}-like function.

Although the Backus–Gilbert method has been designed for linear problems, its extension to nonlinear problems is straightforward. Let us consider the update formula

$$\mathbf{x}^\delta_{k+1} = \mathbf{x}^\delta_k + \mathbf{p}^\delta_k, \quad k = 0, 1, \ldots,$$

where $\mathbf{p}^\delta_k$ is the Newton step and $\mathbf{x}^\delta_0 = \mathbf{x}_\mathtt{a}$. Further, let $\mathbf{p}^\dagger_k = \mathbf{x}^\dagger - \mathbf{x}^\delta_k$ be the exact step, where $\mathbf{x}^\dagger$ is a solution of the nonlinear equation with exact data $\mathbf{F}(\mathbf{x}) = \mathbf{y}$. It is quite obvious that $\mathbf{p}^\dagger_k$ solves the equation (see Appendix H)

$$\mathbf{K}_k \mathbf{p} = \mathbf{r}_k, \tag{9.10}$$

with

$$\mathbf{r}_k = \mathbf{y} - \mathbf{F}\left(\mathbf{x}^\delta_k\right) - \mathbf{R}\left(\mathbf{x}^\dagger, \mathbf{x}^\delta_k\right) \tag{9.11}$$

and $\mathbf{R}\left(\mathbf{x}^\dagger, \mathbf{x}^\delta_k\right)$ being the linearization error. As $\mathbf{r}_k$ is unknown, and only

$$\mathbf{r}^\delta_k = \mathbf{y}^\delta - \mathbf{F}\left(\mathbf{x}^\delta_k\right), \tag{9.12}$$

is available, we consider the equation

$$\mathbf{K}_k \mathbf{p} = \mathbf{r}^\delta_k, \tag{9.13}$$

and compute $\mathbf{p}^\delta_k$ as

$$\mathbf{p}^\delta_k = \mathbf{K}^\dagger_k \mathbf{r}^\delta_k. \tag{9.14}$$

In (9.14), the generalized inverse $\mathbf{K}^\dagger_k$ is unknown and its row vectors will be determined one by one. Before doing this, we observe that the ith entry of $\mathbf{p}^\delta_k$ is given by

$$\left[\mathbf{p}^\delta_k\right]_i = \mathbf{k}^{\dagger T}_i \mathbf{r}^\delta_k, \tag{9.15}$$

where $\mathbf{k}_i^{\dagger T}$ is the ith row vector of $\mathbf{K}_k^{\dagger}$, partitioned as

$$\mathbf{K}_k^{\dagger} = \begin{bmatrix} \mathbf{k}_1^{\dagger T} \\ \vdots \\ \mathbf{k}_n^{\dagger T} \end{bmatrix}. \tag{9.16}$$

Now, defining as usual the averaging kernel matrix $\mathbf{A}_k$ by

$$\mathbf{A}_k = \mathbf{K}_k^{\dagger}\mathbf{K}_k, \tag{9.17}$$

and assuming the partitions

$$\mathbf{A}_k = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix}, \quad \mathbf{K}_k = [\mathbf{k}_1, \ldots, \mathbf{k}_n],$$

we obtain

$$[\mathbf{a}_i]_j = \mathbf{k}_i^{\dagger T}\mathbf{k}_j, \quad i, j = 1, \ldots, n. \tag{9.18}$$

To compute the row vector $\mathbf{k}_i^{\dagger T}$ we proceed to formulate the constrained minimization problem (9.7) in terms of the averaging kernel $\mathbf{a}_i^T$. For this purpose, we discretize the altitude interval $[0, z_{\max}]$ in n layers and put $[\mathbf{a}_i]_j = a_\mu(z_i, z_j)$, where z_i is the centerpoint of the layer i. The objective function in (9.7) can then be expressed as (cf. (9.18))

$$\begin{aligned} s(z_i) &= \int_0^{z_{\max}} d_\mu(z_i, z)\, a_\mu(z_i, z)^2 \, \mathrm{d}z \\ &= \sum_{j=1}^{n} d_\mu(z_i, z_j)\, a_\mu(z_i, z_j)^2 \triangle z_j \\ &= \sum_{j=1}^{n} d_\mu(z_i, z_j)\, [\mathbf{a}_i]_j^2 \triangle z_j \\ &= \mathbf{k}_i^{\dagger T}\mathbf{Q}_{ki}\mathbf{k}_i^{\dagger}, \end{aligned}$$

where $\triangle z_i$ is the geometrical thickness of the layer i, and

$$\mathbf{Q}_{ki} = \mathbf{K}_k \left[\mathrm{diag}\left(d_\mu(z_i, z_j) \triangle z_j\right)_{n\times n}\right] \mathbf{K}_k^T.$$

For the choice (9.8) with $\mu = 2$, $s(z_i)$ represents the spread of the averaging kernel around z_i, and by minimizing the spread we intend to guarantee that the resolution of the retrieval is as high as possible. The normalization condition in (9.7) takes the form (cf. (9.18))

$$1 = \int_0^{z_{\max}} a_\mu(z_i, z)\, \mathrm{d}z = \sum_{j=1}^{n} a_\mu(z_i, z_j) \triangle z_j = \sum_{j=1}^{n} [\mathbf{a}_i]_j \triangle z_j = \mathbf{k}^T\mathbf{k}_i^{\dagger},$$

with

$$\mathbf{k} = \sum_{j=1}^{n} \mathbf{k}_j \triangle z_j,$$

and the constrained minimization problem to be solved reads as

$$\min_{\mathbf{k}^\dagger} \mathbf{k}^{\dagger T}\mathbf{Q}_{ki}\mathbf{k}^\dagger \quad \text{subject to } \mathbf{k}^T\mathbf{k}^\dagger = 1. \tag{9.19}$$

Via the Lagrange multiplier formalism, the row vector $\mathbf{k}_i^{\dagger T}$ is determined by minimizing the Lagrangian function

$$\mathcal{L}\left(\mathbf{k}^\dagger, \lambda\right) = \frac{1}{2}\mathbf{k}^{\dagger T}\mathbf{Q}_{ki}\mathbf{k}^\dagger + \lambda\left(\mathbf{k}^T\mathbf{k}^\dagger - 1\right), \tag{9.20}$$

and the result is

$$\mathbf{k}_i^\dagger = \frac{1}{\mathbf{q}_i^T\mathbf{k}}\mathbf{q}_i. \tag{9.21}$$

with

$$\mathbf{q}_i = \mathbf{Q}_{ki}^{-1}\mathbf{k}.$$

In practice it is necessary to add regularization when the problem (9.19) is solved numerically, due to the ill-conditioning of the matrix $\mathbf{Q}_{ki}$. Neglecting the linearization error $\mathbf{R}\left(\mathbf{x}^\dagger, \mathbf{x}_k^\delta\right)$, the Newton step $\mathbf{p}_k^\delta$ can be expressed as (cf. (9.10)–(9.12) and (9.14))

$$\mathbf{p}_k^\delta = \mathbf{K}_k^\dagger\mathbf{r}_k^\delta = \mathbf{K}_k^\dagger\left(\mathbf{r}_k + \boldsymbol{\delta}\right) = \mathbf{A}_k\mathbf{p}_k^\dagger + \mathbf{K}_k^\dagger\boldsymbol{\delta},$$

and it is apparent that the spread accounts only for the smoothed component $\mathbf{A}_k\mathbf{p}_k^\dagger$ of $\mathbf{p}_k^\delta$. The ith entry of the noise error vector $\mathbf{e}_{\mathrm{n}k}^\delta = -\mathbf{K}_k^\dagger\boldsymbol{\delta}$ is

$$\left[\mathbf{e}_{\mathrm{n}k}^\delta\right]_i = -\mathbf{k}_i^{\dagger T}\boldsymbol{\delta},$$

and for white noise with covariance $\mathbf{C}_\delta = \sigma^2\mathbf{I}_m$, the expected value of the noise error is given by

$$n\left(z_i\right) = \mathcal{E}\left\{\left[\mathbf{e}_{\mathrm{n}k}^\delta\right]_i^2\right\} = \sigma^2\left\|\mathbf{k}_i^\dagger\right\|^2. \tag{9.22}$$

In this regard, we construct an objective function reflecting a trade-off between spread and noise error, that is, we consider the constrained minimization problem

$$\min_{\mathbf{k}^\dagger}\left(\mathbf{k}^{\dagger T}\mathbf{Q}_{ki}\mathbf{k}^\dagger + \alpha\left\|\mathbf{k}^\dagger\right\|^2\right) \quad \text{subject to } \mathbf{k}^T\mathbf{k}^\dagger = 1. \tag{9.23}$$

The objective function in (9.23) is as in (9.19), but with $\mathbf{Q}_{ki} + \alpha\mathbf{I}_m$ in place of $\mathbf{Q}_{ki}$; the solution of (9.23) is then

$$\mathbf{k}_{\alpha i}^\dagger = \frac{1}{\mathbf{q}_{\alpha i}^T\mathbf{k}}\mathbf{q}_{\alpha i}, \tag{9.24}$$

with

$$\mathbf{q}_{\alpha i} = \left(\mathbf{Q}_{ki} + \alpha\mathbf{I}_m\right)^{-1}\mathbf{k}. \tag{9.25}$$

Once the row vectors of the generalized inverse have been computed, the Backus–Gilbert step is determined via (cf. (9.15))

$$\left[\mathbf{p}_{k\alpha}^{\delta}\right]_i = \mathbf{k}_{\alpha i}^{\dagger T}\mathbf{r}_k^{\delta} = \frac{\mathbf{q}_{\alpha i}^T\mathbf{r}_k^{\delta}}{\mathbf{q}_{\alpha i}^T\mathbf{k}},\ \ i = 1, \ldots, n. \tag{9.26}$$

Let us discuss some practical implementation issues by following the analysis of Hansen (1994). Defining the diagonal matrices

$$\mathbf{D}_i = \left[\operatorname{diag}\left(\sqrt{d_\mu\left(z_i, z_j\right)}\right)_{n\times n}\right],\ \ \mathbf{Z} = \left[\operatorname{diag}\left(\triangle z_j\right)_{n\times n}\right],$$

and denoting by $\mathbf{e}$ the n-dimensional vector of all ones, i.e., $\mathbf{e} = [1, \ldots, 1]^T$, we express $\mathbf{Q}_{ki}$ as

$$\mathbf{Q}_{ki} = \mathbf{K}_k\mathbf{D}_i\mathbf{Z}\mathbf{D}_i\mathbf{K}_k^T.$$

Setting

$$\bar{\mathbf{K}}_{ki} = \mathbf{K}_k\mathbf{D}_i\mathbf{Z}^{\frac{1}{2}},\ \ \mathbf{e}_i = \mathbf{D}_i^{-1}\mathbf{Z}^{\frac{1}{2}}\mathbf{e},$$

and noting that

$$\mathbf{k} = \mathbf{K}_k\mathbf{Z}\mathbf{e} = \bar{\mathbf{K}}_{ki}\mathbf{e}_i,$$

we write $\mathbf{q}_{\alpha i}$ as (cf. (9.25))

$$\mathbf{q}_{\alpha i} = \left(\bar{\mathbf{K}}_{ki}\bar{\mathbf{K}}_{ki}^T + \alpha\mathbf{I}_m\right)^{-1}\bar{\mathbf{K}}_{ki}\mathbf{e}_i. \tag{9.27}$$

Moreover, we have (cf. (9.24))

$$\mathbf{k}_{\alpha i}^{\dagger} = \frac{1}{\mathbf{q}_{\alpha i}^T\bar{\mathbf{K}}_{ki}\mathbf{e}_i}\mathbf{q}_{\alpha i}$$

and (cf. (9.26))

$$\left[\mathbf{p}_{k\alpha}^{\delta}\right]_i = \frac{\mathbf{q}_{\alpha i}^T\mathbf{r}_k^{\delta}}{\mathbf{q}_{\alpha i}^T\bar{\mathbf{K}}_{ki}\mathbf{e}_i}. \tag{9.28}$$

Note that the singularity of $\mathbf{D}_i^{-1}$ at $j = i$ can be removed in practice by approximating

$$d_\mu\left(z_i, z_i\right) \approx d_\mu\left(z_i, z_i + \triangle z\right),$$

with $\triangle z$ sufficiently small, e.g., $\triangle z = 1$ m. An inspection of (9.27) reveals that $\mathbf{q}_{\alpha i}$ minimizes the Tikhonov function

$$\mathcal{F}_\alpha\left(\mathbf{q}\right) = \left\|\mathbf{e}_i - \bar{\mathbf{K}}_{ki}^T\mathbf{q}\right\|^2 + \alpha\left\|\mathbf{q}\right\|^2.$$

Thus, if $(\bar{\sigma}_j; \bar{\mathbf{v}}_j, \bar{\mathbf{u}}_j)$ is a singular system of $\bar{\mathbf{K}}_{ki}$, we obtain the representation

$$\mathbf{q}_{\alpha i} = \sum_{j=1}^{n}\frac{\bar{\sigma}_j}{\bar{\sigma}_j^2 + \alpha}\left(\bar{\mathbf{v}}_j^T\mathbf{e}_i\right)\bar{\mathbf{u}}_j,$$

and the useful expansions

$$\mathbf{q}_{\alpha i}^T \bar{\mathbf{K}}_{ki} \mathbf{e}_i = \sum_{j=1}^{n} \frac{\bar{\sigma}_j^2}{\bar{\sigma}_j^2 + \alpha} \left(\bar{\mathbf{v}}_j^T \mathbf{e}_i \right)^2 , \tag{9.29}$$

and

$$\mathbf{q}_{\alpha i}^T \mathbf{r}_k^\delta = \sum_{j=1}^{n} \frac{\bar{\sigma}_j^2}{\bar{\sigma}_j^2 + \alpha} \frac{1}{\bar{\sigma}_j} \left(\bar{\mathbf{v}}_j^T \mathbf{e}_i \right) \left(\bar{\mathbf{u}}_j^T \mathbf{r}_k^\delta \right) . \tag{9.30}$$

The Backus–Gilbert solution can be computed for any value of the regularization parameter α, by inserting (9.29) and (9.30) into (9.28).

To reveal the regularizing effect of the Backus–Gilbert method we mention that the characteristic features of the singular vectors of $\mathbf{K}_k$ carry over to the singular vectors of $\bar{\mathbf{K}}_{ki}$, and that the filter factors in (9.30) damp out the noisy components in the data as the Tikhonov filter factors do.

To compute the regularization parameter we may impose that the noise error (9.22) has a prescribed value, that is,

$$n_\alpha \left(z_i \right) = \varepsilon_{\mathtt{n}} \left[\mathbf{x}_{\mathtt{a}} \right]_i^2 , \tag{9.31}$$

for some relative error level $\varepsilon_{\mathtt{n}}$. Another selection criterion can be designed by taking into account that the spread is an increasing function of α and that the noise error is a decreasing function of α. Thus, we may follow the idea of the L-curve method, and compute the regularization parameter which balances the spread and noise error. For any value of α, the computable expressions of the quantities of interest are

$$s_\alpha \left(z_i \right) = \frac{1}{\left(\mathbf{q}_{\alpha i}^T \bar{\mathbf{K}}_{ki} \mathbf{e}_i \right)^2} \sum_{j=1}^{n} \left(\frac{\bar{\sigma}_j^2}{\bar{\sigma}_j^2 + \alpha} \bar{\mathbf{v}}_j^T \mathbf{e}_i \right)^2 ,$$

$$n_\alpha \left(z_i \right) = \frac{\sigma^2}{\left(\mathbf{q}_{\alpha i}^T \bar{\mathbf{K}}_{ki} \mathbf{e}_i \right)^2} \sum_{j=1}^{n} \left(\frac{\bar{\sigma}_j}{\bar{\sigma}_j^2 + \alpha} \bar{\mathbf{v}}_j^T \mathbf{e}_i \right)^2 ,$$

and the regularization parameter, corresponding to the point on the curve at which the tangent has the slope -1, is chosen as the minimizer of the function (Reginska, 1996),

$$\beta \left(\alpha \right) = x \left(\alpha \right) + y \left(\alpha \right) , \tag{9.32}$$

with $x \left(\alpha \right) = s_\alpha$ and $y \left(\alpha \right) = n_\alpha$.

In Figure 9.1 we plot the solution errors for the O_3 retrieval test problem. The δ^{-1}-like functions (9.8) and (9.9) yield similar accuracies but for different domains of variation of the regularization parameter. The regularizing effect of the Backus–Gilbert method is also apparent in this figure: by increasing the signal-to-noise ratio, the minimum solution error as well as the optimal value of the regularization parameter (the minimizer) decrease.

In our numerical analysis we used a discrete version of the regularization parameter choice methods (9.31) and (9.32), that is, for the set $\{\alpha_j\}$ with $\alpha_j = \bar{\sigma}_j^2$, $j = 1, \ldots, n$, we chose the regularization parameter $\alpha_{j^\star}$ as the smallest α_j satisfying $n_{\alpha_{j^\star}} \left(z_i \right) \leq \varepsilon_{\mathtt{n}} \left[\mathbf{x}_{\mathtt{a}} \right]_i^2$, or as the minimizer of $\beta \left(\alpha_j \right)$. The plots in Figure 9.2 illustrate that the noise error is

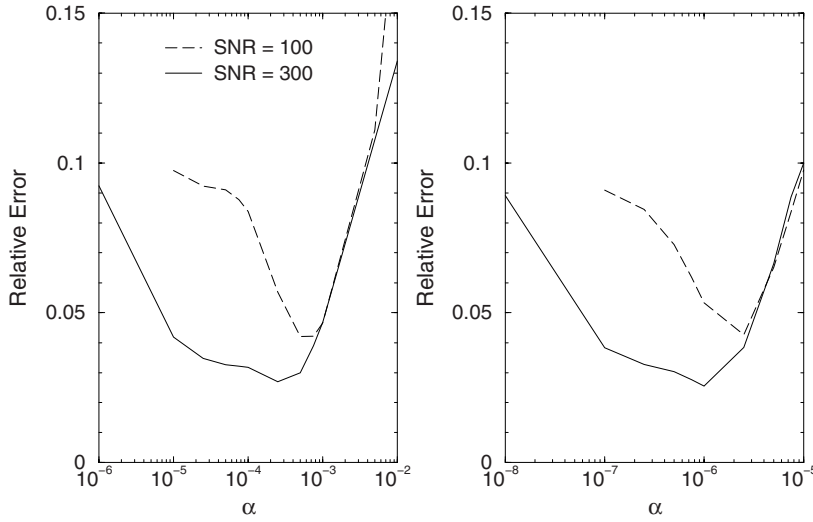

Fig. 9.1. Relative solution errors for the Backus–Gilbert method with the quadratic function (9.8) (left) and the exponential function (9.9) (right). The parameters of calculation are $\mu = 2$ and $l = 1.0$ km for the quadratic function, and $\mu = 2$ and $l = 10.0$ km for the exponential function.

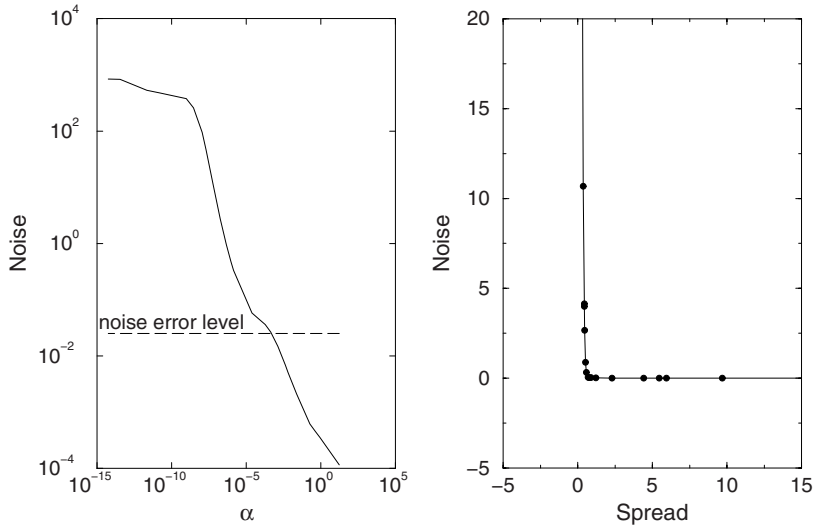

Fig. 9.2. Noise error curve (left) and L-curve (right) for a layer situated at 30.6 km.

a decreasing function of the regularization parameter and that the L-curve has a distinct corner.

The solution errors given in Table 9.1 show that the noise error criterion yields sufficiently accurate results. By contrast, the L-curve method predicts a value of the regularization parameter which is considerably smaller than the optimal value. As a result,

Table 9.1. Relative solution errors for the Backus–Gilbert method with the noise error (NE) criterion and the L-curve (LC) method.

δ^{-1}-like function	SNR	Method	ε	$\varepsilon_{\mathtt{opt}}$
quadratic	100	NE	6.65e-2	4.21e-2
		LC	2.30e-1	
	300	NE	5.74e-2	2.51e-2
		LC	1.42e-1	
exponential	100	NE	5.88e-2	4.26e-2
		LC	3.15e-1	
	300	NE	5.41e-2	2.55e-2
		LC	2.42e-1	

the retrieved profiles are undersmoothed (Figure 9.3). Note that the failure of the L-curve method is because we use a very rough discrete search procedure to minimize β.

In the framework of mollifier methods, the approximate generalized inverse is determined independently of the data, and therefore, mollifier methods can be viewed as being equivalent to Tikhonov regularization with an a priori parameter choice method. In practice, the methods are computationally very expensive because for each layer, we have to solve an optimization problem. However, for the operational usage of a near real-time software processor, this drawback is only apparent; when the approximate generalized inverse

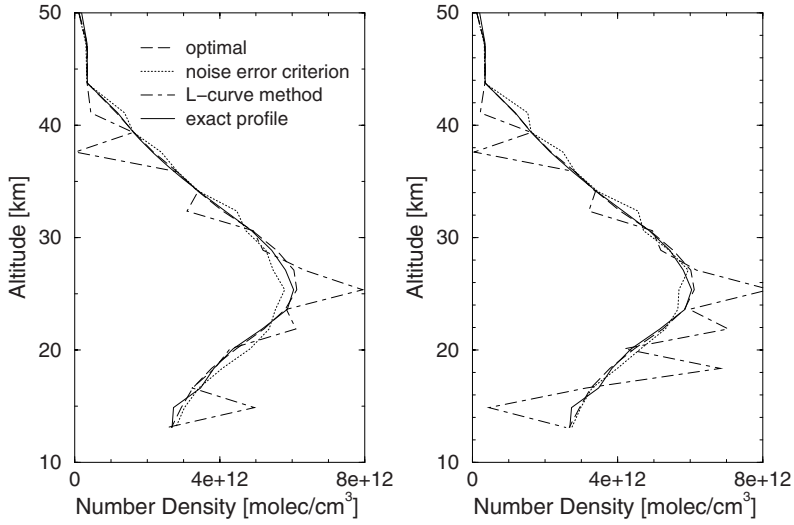

Fig. 9.3. Retrieved profiles computed with the Backus–Gilbert method using the quadratic function (9.8) (left) and the exponential function (9.9) (right). The curves correspond to the optimal value of the regularization parameter (the minimizer in Figure 9.1), the noise error criterion and the L-curve method. The signal-to-noise ratio is SNR = 300 and the parameters of calculation are as in Figure 9.1.

is (a priori) computed and stored, the processing of data is much faster than, for example, Tikhonov regularization with an a posteriori parameter choice method, because it involves only matrix-vector multiplications.

9.2 Maximum entropy regularization

First proposed as a general inference procedure by Jaynes (1957) on the basis of Shannon's axiomatic characterization of the amount of information (Shannon, 1949; Shannon and Weaver, 1949), the maximum entropy principle emerged as a successful regularization technique due to the contributions of Frieden (1972), and Gull and Daniel (1978). Although the conventional maximum entropy regularization operates with the concept of absolute entropy (or Shannon entropy), we describe a formulation based on relative and cross entropies, which allows a better exploitation of the available a priori information (Engl et al., 2000).

To sketch the maximum entropy regularization we consider a discrete random variable X with a finite number of realizations $x_1, \ldots, x_n$, and suppose that we make some a priori assumptions about the probability mass function of X,

$$p_{\mathtt{a}}(x) = \begin{cases} p_{\mathtt{a}i}, & X = x_i, \\ 0, & \text{otherwise}, \end{cases} \qquad \sum_{i=1}^{n} p_{\mathtt{a}i} = 1.$$

By measurements we obtain additional information on X, which lets us change our a priori probability mass function into the a posteriori probability mass function,

$$p(x) = \begin{cases} p_i, & X = x_i, \\ 0, & \text{otherwise}, \end{cases} \qquad \sum_{i=1}^{n} p_i = 1.$$

We recall that in statistical inversion theory, the a posteriori probability mass function represents the conditional probability density of X given the measurement data. The goal of our analysis is the computation of the a posteriori probability mass function by considering the new data.

In information theory, a natural distance measure from the probability mass function p to the probability mass function $p_{\mathtt{a}}$ is the Kullback–Leibler divergence defined by

$$D(p; p_{\mathtt{a}}) = \sum_{i=1}^{n} p_i \log\left(\frac{p_i}{p_{\mathtt{a}i}}\right).$$

Essentially, the Kullback–Leibler divergence signifies the amount of useful information about X, that can be obtained given the measurements. The negative of the Kullback–Leibler divergence represents the relative entropy

$$H_{\mathtt{r}}(p; p_{\mathtt{a}}) = -\sum_{i=1}^{n} p_i \log\left(\frac{p_i}{p_{\mathtt{a}i}}\right).$$

Note that as opposed to the absolute entropy $H(p) = -\sum_{i=1}^{n} p_i \log p_i$, the relative entropy $H_{\mathtt{r}}$ is negative (cf. (9.35) below) and attains its global maximum $H_{\mathtt{rmax}} = 0$ at $p = p_{\mathtt{a}}$.

To compute the a posteriori probability mass function, we minimize the Kullback–Leibler divergence D (or maximize the relative entropy $H_{\mathtt{r}}$) with the data and the normalization condition $\sum_{i=1}^{n} p_i = 1$ as constraints. If $\mathbf{x}$ is the state vector to be retrieved and $\mathbf{x}_{\mathtt{a}}$ is the a priori state, we define the normalized vectors

$$\bar{\mathbf{x}} = \frac{1}{\sum_{i=1}^{n} [\mathbf{x}]_i}\mathbf{x}, \quad \bar{\mathbf{x}}_{\mathtt{a}} = \frac{1}{\sum_{i=1}^{n} [\mathbf{x}_{\mathtt{a}}]_i}\mathbf{x}_{\mathtt{a}},$$

and under the assumptions $[\mathbf{x}]_i > 0$ and $[\mathbf{x}_{\mathtt{a}}]_i > 0$ for $i = 1, \ldots, n$, we interpret the components of these vectors as the probabilities p_i and $p_{\mathtt{a}i}$, respectively. As data we consider the nonlinear model $\mathbf{y}^{\delta} = \mathbf{F}(\mathbf{x}) + \boldsymbol{\delta}$, and impose the feasibility constraint

$$\left\|\mathbf{y}^{\delta} - \mathbf{F}(\mathbf{x})\right\|^2 \leq \Delta^2.$$

The constrained minimization problem then takes the form

$$\begin{aligned} &\min_{\mathbf{x}} \Lambda_{\mathtt{r}}(\mathbf{x}) = \sum_{i=1}^{n} [\bar{\mathbf{x}}]_i \log\left(\frac{[\bar{\mathbf{x}}]_i}{[\bar{\mathbf{x}}_{\mathtt{a}}]_i}\right) \\ &\text{subject to } \left\|\mathbf{y}^{\delta} - \mathbf{F}(\mathbf{x})\right\|^2 \leq \Delta^2. \end{aligned} \tag{9.33}$$

By virtue of the Lagrange multiplier formalism, the problem (9.33) is equivalent to the minimization of the Tikhonov function

$$\mathcal{F}_{\alpha}(\mathbf{x}) = \frac{1}{2}\left\|\mathbf{y}^{\delta} - \mathbf{F}(\mathbf{x})\right\|^2 + \alpha\Lambda_{\mathtt{r}}(\mathbf{x}). \tag{9.34}$$

Using the inequality

$$\log z \geq 1 - \frac{1}{z}, \quad z > 0, \tag{9.35}$$

we find that

$$\Lambda_{\mathtt{r}}(\mathbf{x}) \geq \sum_{i=1}^{n} ([\bar{\mathbf{x}}]_i - [\bar{\mathbf{x}}_{\mathtt{a}}]_i) = 0.$$

Evidently, the global minimizer of $\Lambda_{\mathtt{r}}$ is attained for $\bar{\mathbf{x}} = \bar{\mathbf{x}}_{\mathtt{a}}$, which reiterates the role of $\mathbf{x}_{\mathtt{a}}$ as a priori information.

If $\mathbf{x}$ and $\mathbf{x}_{\mathtt{a}}$ are not normalized, the non-negative functions (Eggermont, 1993)

$$\Lambda_{\mathtt{B}}(\mathbf{x}) = \sum_{i=1}^{n} \left[\log\left(\frac{[\mathbf{x}]_i}{[\mathbf{x}_{\mathtt{a}}]_i}\right) + \frac{[\mathbf{x}_{\mathtt{a}}]_i}{[\mathbf{x}]_i} - 1\right]$$

and

$$\Lambda_{\mathtt{c}}(\mathbf{x}) = \sum_{i=1}^{n} \left[\frac{[\mathbf{x}]_i}{[\mathbf{x}_{\mathtt{a}}]_i}\log\left(\frac{[\mathbf{x}]_i}{[\mathbf{x}_{\mathtt{a}}]_i}\right) - \frac{[\mathbf{x}]_i}{[\mathbf{x}_{\mathtt{a}}]_i} + 1\right],$$

representing the negative of the Burg's entropy and the cross entropy, respectively, can be used as penalty terms. A Taylor expansion of the cross entropy about the a priori yields

$$\Lambda_{\mathtt{c}}(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_{\mathtt{a}})^T \left[\operatorname{diag}\left(\frac{1}{[\mathbf{x}_{\mathtt{a}}]_i^2}\right)_{n\times n}\right](\mathbf{x} - \mathbf{x}_{\mathtt{a}}) + O\left(\|\mathbf{x} - \mathbf{x}_{\mathtt{a}}\|^3\right),$$

and we see that in the neighborhood of the a priori, the cross entropy regularization matrix behaves like a diagonal matrix.

Ramos et al. (1999), following the work of Landl and Anderson (1996), developed two entropic regularization techniques by using penalty functions which are similar to the discrete difference operators. The first-order penalty function (corresponding to the entropy of the vector of first-order differences of $\mathbf{x}$) is defined by

$$\Lambda_1(\mathbf{x}) = \sum_{i=1}^{n-1} \frac{(n-1)\, d_{1i}}{\sum_{i=1}^{n-1} d_{1i}} \log\left(\frac{(n-1)\, d_{1i}}{\sum_{i=1}^{n-1} d_{1i}} \right),$$

where the d_{1i} can be chosen as

$$d_{1i} = \left([\mathbf{x}]_{i+1} - [\mathbf{x}]_i \right) + (x_{\max} - x_{\min}) + \varsigma, \;\; i = 1, \ldots, n-1, \tag{9.36}$$

or as

$$d_{1i} = \left| [\mathbf{x}]_{i+1} - [\mathbf{x}]_i \right| + \varsigma, \;\; i = 1, \ldots, n-1. \tag{9.37}$$

Here, ς is a small positive constant, while $x_{\min}$ and $x_{\max}$ are the lower and the upper bounds of all entries in $\mathbf{x}$, that is, and $x_{\min} \leq [\mathbf{x}]_i \leq x_{\max}$, $i = 1, \ldots, n$. By (9.35), we have

$$\frac{(n-1)\, d_{1i}}{d_1} \log\left(\frac{(n-1)\, d_{1i}}{d_1} \right) \geq \frac{(n-1)\, d_{1i}}{d_1} - 1,$$

with $d_1 = \sum_{i=1}^{n-1} d_{1i}$, and we infer that $\Lambda_1 \geq 0$. The minimum value of Λ_1 is attained when all d_{1i} are the same, and the solutions to (9.34) approach the discrete approximation of a first-order polynomial as $\alpha \to \infty$. The second-order penalty function (corresponding to the entropy of the vector of second-order differences of $\mathbf{x}$) is given by

$$\Lambda_2(\mathbf{x}) = \sum_{i=2}^{n-1} \frac{(n-2)\, d_{2i}}{\sum_{i=2}^{n-1} d_{2i}} \log\left(\frac{(n-2)\, d_{2i}}{\sum_{i=2}^{n-1} d_{2i}} \right),$$

with

$$d_{2i} = \left([\mathbf{x}]_{i+1} - 2[\mathbf{x}]_i + [\mathbf{x}]_{i-1} \right) + 2(x_{\max} - x_{\min}) + \varsigma, \;\; i = 2, \ldots, n-1, \tag{9.38}$$

or

$$d_{2i} = \left| [\mathbf{x}]_{i+1} - 2[\mathbf{x}]_i + [\mathbf{x}]_{i-1} \right| + \varsigma, \;\; i = 2, \ldots, n-1. \tag{9.39}$$

As before, $\Lambda_2 \geq 0$ attains its minimum when all d_{2i} coincide, and the solutions to (9.34) approach the discrete approximation of a second-order polynomial as $\alpha \to \infty$. In comparison, under similar conditions, Tikhonov regularization with the first- and second-order difference regularization matrices will yield a constant solution and a straight line, respectively.

The minimization of the Tikhonov function (9.34) can be performed by using the Newton method with

$$\mathbf{g}_\alpha(\mathbf{x}) = \nabla \mathcal{F}_\alpha(\mathbf{x}) = \mathbf{K}(\mathbf{x})^T \left[\mathbf{F}(\mathbf{x}) - \mathbf{y}^\delta \right] + \alpha \nabla \Lambda(\mathbf{x}),$$

and the Hessian approximation

$$\mathbf{G}_\alpha(\mathbf{x}) = \nabla^2 \mathcal{F}_\alpha(\mathbf{x}) \approx \mathbf{K}(\mathbf{x})^T \mathbf{K}(\mathbf{x}) + \alpha \nabla^2 \Lambda(\mathbf{x}).$$

To be more concrete, at the iteration step k, the search direction $\mathbf{p}_{\alpha k}^\delta$ is computed as the solution of the Newton equation

$$\mathbf{G}_\alpha\left(\mathbf{x}_{\alpha k}^\delta\right)\mathbf{p} = -\mathbf{g}_\alpha\left(\mathbf{x}_{\alpha k}^\delta\right),$$

the step length τ_k is determined by imposing the descent condition, and the new iterate is taken as $\mathbf{x}_{\alpha k+1}^\delta = \mathbf{x}_{\alpha k}^\delta + \tau_k \mathbf{p}_{\alpha k}^\delta$.

In Figure 9.4 we plot the retrieved O_3 profiles for the cross entropy regularization with the penalty term Λ_c. Because in this case, the regularization matrix acts like a diagonal matrix, the solution errors may become extremely large. Specifically, on a fine grid, the number densities, with respect to which the retrieval is insensitive, are close to the a priori.

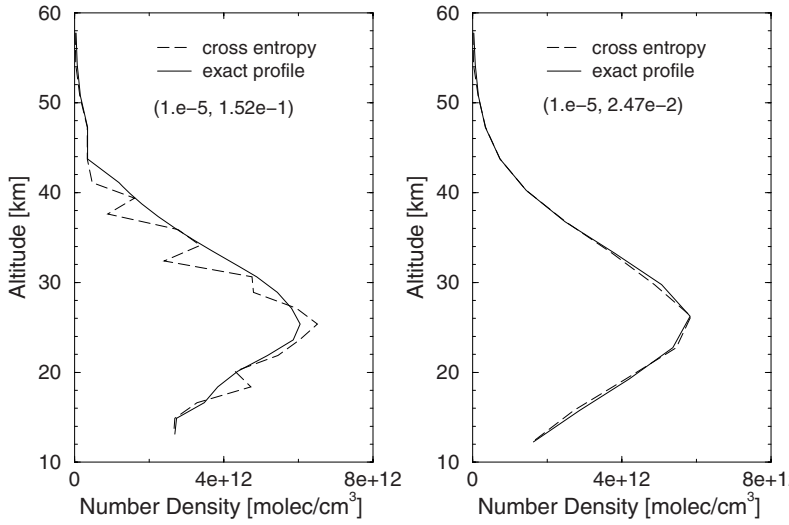

Fig. 9.4. Retrieved O_3 profiles computed with the cross entropy regularization on a retrieval grid with 36 levels (left) and on a retrieval grid with 24 levels (right). The numbers in parentheses represent the values of the regularization parameter and of the relative solution error.

The plots in Figure 9.5 illustrate the solution errors for the first- and second-order entropy regularization with the penalty terms Λ_1 and Λ_2, respectively. As for Tikhonov regularization, the error curves possess a minimum for an optimal value of the regularization parameter. The minima of the solution errors are $3.32 \cdot 10^{-2}$ and $5.05 \cdot 10^{-2}$ for the first-order entropy regularization with the selection criteria (9.36) and (9.37), respectively, and $3.79 \cdot 10^{-2}$ and $2.73 \cdot 10^{-2}$ for the second-order entropy regularization with the selection criteria (9.38) and (9.39), respectively. Comparing both regularization methods we observe that

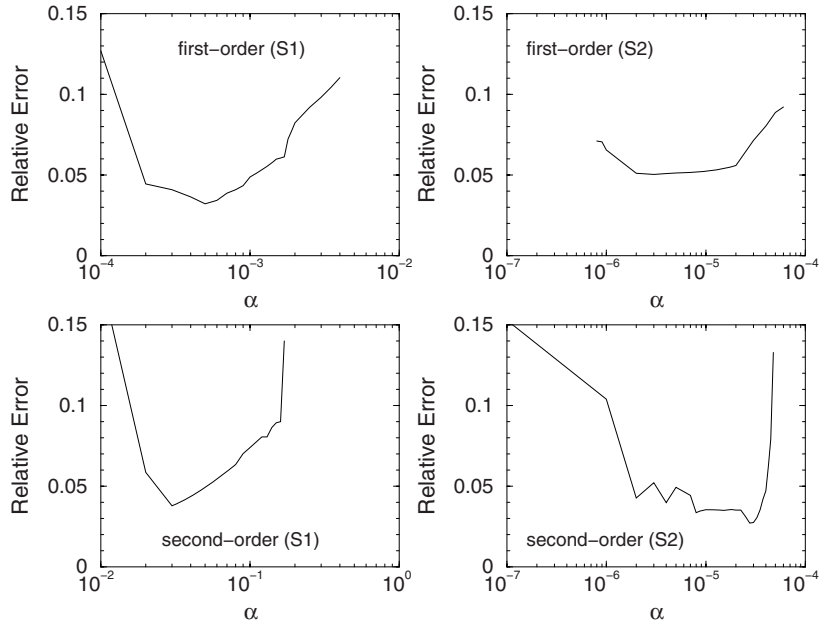

Fig. 9.5. Relative solution errors for the first-order entropy regularizations with the selection criteria (9.36) (S1) and (9.37) (S2), and the second-order entropy regularization with the selection criteria (9.38) (S1) and (9.39) (S2).

(1) the first- and second-order entropy regularizations yield results of comparable accuracies;
(2) the domains of variation of the regularization parameter with acceptable reconstruction errors are larger for the selection criteria (9.37) and (9.39).

A pertinent analysis of the maximum entropy regularization can be found in Engl et al. (2000), while for applications of the second-order entropy regularization in atmospheric remote sensing we refer to Steinwagner et al. (2006).