12. Algorithms for Optimal Smoothing Parameter

12.1. Introduction

As before, let the linear continuous operators $A : X \to Z$, $\quad T : X \to Y$ be defined in the Hilbert spaces, z be an element of the space Z. Present as in Chapter 1, the variational principle for the interpolating spline $\sigma \in X$ in the following way

$$\sigma = \arg \min_{u \in X,\, Au=z} \|Tu\|_Y \tag{12.1}$$

and for the smoothing spline $\sigma_\alpha \in X$ with $\alpha > 0$ -

$$\sigma_\alpha = \arg \min_{u \in X} \alpha \|Tu\|_Y^2 + \|Au - z\|_Z^2. \tag{12.2}$$

A more practical problem is the problem of spline approximation in the convex set

$$\sigma^\varepsilon = \arg \min_{u \in X,\, \|Au - z\|_Z \le \varepsilon} \|Tu\|_Y, \tag{12.3}$$

where as opposed to the interpolating spline σ, the element σ^ε is sought for, whose image $A\sigma^\varepsilon$ does not coincide with the element $z \in Z$, but lies only in its ε-neighbourhood. Problem (12.3) is useful in practice due to various reasons, among which we distinguish:

1) The element z may be known with an error, which must be taken into account;

2) The element z may not lie in the range of the operator A, and the interpolating spline probably does not exist unlike the spline in the convex set.

Very interesting relations take place beetween the splines σ, σ_α and σ^ε. They turn out to be elements of the same space Sp(T, A). Moreover, the space of interpolating splines for different $z \in Z$, the space of smoothing splines for different $z \in Z$, $\alpha > 0$ and the space of splines in convex sets for different $z \in Z$, $\varepsilon > 0$ coincide. In Chapter 1, it was noted that "any smoothing spline is some interpolating". This may be continued in different variations, like "any spline on a convex set is some smoothing spline", and so on. Remember two theorems about splines: the first is from Chapter 1 and the second is from (Laurent 1970, et al.).

Theorem 12.1. If the kernels of the operators T and A - $N(T)$ and $N(A)$ cross only on zero element, and the linear set $N(T) + N(A)$ is closed, then the solution to problem (12.2) is uniquely solvable. If, in addition, the sets $Au = z$ and $\|Au - z\|_Z \leq \varepsilon$ are non empty, then the solutions to problems (12.1) and (12.3) are uniquely solvable, respectively.

Theorem 12.2. Define the residual function

$$\varphi(\alpha) = \|A\sigma_\alpha - z\|_Z \tag{12.4}$$

and the parameters

$$\varepsilon_{\min} = \min_{u \in X} \|Au - z\|, \quad \varepsilon_{\max} = \min_{u \in N(T)} \|Au - z\|.$$

If the conditions of uniqueness in the previous theorem are fulfilled, then

$$\sigma^\varepsilon = \sigma_{\varphi^{-1}(\varepsilon)}, \quad \forall \varepsilon_{\min} < \varepsilon < \varepsilon_{\max}, \tag{12.5}$$

$$\|A\sigma^\varepsilon - z\|_Z = \varepsilon. \tag{12.6}$$

Equality (12.5) signifies that any spline on the convex set is some smoothing spline for some element $z \in Z$ and for an unknown parameter $\alpha > 0$ and the function $\varphi(\alpha)$ of the correspondence beetween the parameters α and ε is one-to-one mapping of the positive axis $(0, \infty)$ onto the interval $(\varepsilon_{\min}, \varepsilon_{\max})$. Equality (12.6) follows from (12.5) and affirms that the solution to problem (12.3) lies on the bound of the admissible set $\|Au - z\|_Z \leq \varepsilon$. Thus, problem (12.3) is reduced to simpler problem (12.2), if the suitable smoothing parameter α is given.

The suitable parameter can be obtained from the equation $\varphi(\alpha) = \varepsilon$. The latter one is nonlinear and to this solve one applies the Newton method. To accelarate the rate of convergence one transforms the equation to the equivalent form:

$$\psi(p) = \varepsilon^{-1}, \tag{12.7}$$

where $\psi(p) = \varphi^{-1}(1/p), \quad \alpha = 1/p$.

In Sect. 12.2 we write out the implicit expressions for smoothing spline-operators, for smoothing splines and for the residual function with the help of spectral functions. We generalize the proof of convexity of the function $\psi(p)$, which formerly was know in particular cases. The property of convexity is a necessary grounds for the convergence of the Newton method.

The main sense of the next sections is to present new algorithms of the search for the smoothing parameter by the known residual ε, more exactly, to expand the area of the known algorithms in application to a smoothing problem. First, together with the Newton method we propose the Chebyshev method of the third degree. It is reasonable due to the higher rates of convergence, simplicity of realization formulas, and some other reasons.

Secondly, we propose to begin iterations for these method with the value $p = 0$, that, incidentally, was suggested earlier in (Gordonova, Morozov 1973), but in practice (LIDA-3 1987) it was not used, because nobody could calculate derivatives of the function $\psi(p)$ with $p = 0$. In this Chapter we show how to do it in the abstract form and for two most practic methods of spline-smoothing. The latter ones based on reproducing mappings and on finite-dimensional approach.

Thirdly, on the basis of the proved facts about the opportunity of extrapolation of a spline on smoothing parameter with the help of the Taylor expansions we propose a reccurrent algorithm of approximate calculation of the smoothing spline and its derivatives. This allows us to modify methods in such a way that the construction of the interpolating spline or the spline on a convex set will be equivalent to solving a few smoothing problems with the same smoothing parameter α, but with different elements $z_1, z_2, \ldots$ from Z.

The advantage of the latter way in practice consists in the fact that one needs to solve the same linear system of algebraic equations with different right-hands sides. This allows one to decompose the matrix of the system only one time, and then to resolve the system for various right-hand sides with low costs. In the conventional way, the matrix of the system is altered on iterations.

12.2. Spectral Decomposition of Operators for Smoothing Spline Problem

Setting $p = 1/\alpha$ and $\Sigma_p = \sigma_\alpha$, we turn to the following variational problem

$$\Sigma_p = \arg\min_{u \in X} \|Tu\|_Y^2 + p\|Au - z\|_Z^2,$$

which is equivalent to (12.2). According to Chapter 1, the solution to the problem Σ_p is determined from operator equation

$$(T^*T + pA^*A)\Sigma_p = pA^*z, \tag{12.8}$$

which is uniquely defined in the conditions of Theorem 12.1 for each $p > 0$. Besides, the symmetric and bounded operator $T^*T + A^*A$ becomes positively defined, consequently, one can take its square root $S = \sqrt{T^*T + A^*A}$, which is also symmetric, bounded and positively defined. Clearly, the operator $S^{-1}A^*AS^{-1}$ is symmetric and its spectrum belongs to the interval $[0, 1]$. The Hilbert theorem (Riesz, SZ.-Nagy 1972) asserts that such an operator admits decomposition

$$S^{-1}A^*AS^{-1} = \int_0^1 \lambda dE_\lambda, \tag{12.9}$$

where $\{E_\lambda\}$ is a family of projection operators.

Since $S^{-1}A^*AS^{-1} + S^{-1}T^*TS^{-1} = E$, then

$$S^{-1}T^*TS^{-1} = \int_0^1 (1-\lambda)dE_\lambda. \tag{12.10}$$

On the basis of equalities (12.9-12.10) and properties of the spectral representation of operators, write out the following decompositions of the operators, related to the smoothing spline problem,

$$A^*A = S\int_0^1 \lambda dE_\lambda S, \quad T^*T = S\int_0^1 (1-\lambda)dE_\lambda S,$$

$$T^*T + pA^*A = S\int_0^1 (\lambda p + (1-\lambda))dE_\lambda S,$$

$$(T^*T + pA^*A)^{-1} = S^{-1}\int_0^1 \frac{1}{\lambda p + (1-\lambda)}dE_\lambda S^{-1},$$

$$(T^*T + pA^*A)^{-1}pA^*A = S^{-1}\int_0^1 \frac{\lambda p}{\lambda p + (1-\lambda)}dE_\lambda S,$$

$$A^*A(T^*T + pA^*A)^{-1}pA^*A = S\int_0^1 \frac{\lambda^2 p}{\lambda p + (1-\lambda)}dE_\lambda S.$$

The equation $A^*Ar = A^*z$ may be always solved. So let us solve equation (12.8) in the following way

$$\Sigma_p = (T^*T + pA^*A)^{-1}pA^*Ar.$$

Introduce the residual function $\varphi(p) = \|A(\Sigma_p - r)\|_Z^2$, which slightly differs from the conventional form $\varphi(p) = \|A\Sigma_p - z\|_Z^2$ and coincides with it in the case $Ar = z$. The latter is valid, if the interpolating problem is uniquely defined. For the sake of the future objectives, we write out the residual function in the equivalent form

$$\varphi(p) = (\Sigma_p - r, A^*A(\Sigma_p - r)).$$

Owing to the above-presented spectral decompositions of operators for the smoothing spline problem, the factors of the latter scalar product may be writen out in the following form:

$$\Sigma_p - r = [(T^*T + pA^*A)^{-1}pA^*A - I]r = S^{-1}\int_0^1 \frac{\lambda - 1}{\lambda p + (1-\lambda)}dE_\lambda Sr,$$

$$A^*A(\Sigma_p - r) = A^*A[(T^*T + pA^*A)^{-1}pA^*A - I]r$$
$$= S\int_0^1 \frac{\lambda(\lambda-1)}{\lambda p + (1-\lambda)} dE_\lambda Sr.$$

Multiplying the factors we have:

$$\varphi(p) = \int_0^1 \frac{\lambda(\lambda-1)^2}{(\lambda p + (1-\lambda))^2} d(E_\lambda Sr, Sr),$$

and, finally, taking into account the fact that E_λ is a projection operator,

$$\varphi(p) = \int_0^1 \frac{\lambda(\lambda-1)^2}{(\lambda p + (1-\lambda))^2} d\|E_\lambda Sr\|^2.$$

On the basis of this equality it is easy to prove the subsequent theorem, which substantiates the applicability of the Newton method for the algorithms of the sought for optimal smoothing parameter. Earlier, this theorem was proved independenly for a finite-dimensional case by Reinsch and Morozov.

Theorem 12.3. The function $\psi(p) = \varphi^{-1/2}(p)$ is an increased concave function for $p \geq 0$.

Proof. To investigate isogeometrical properties of the function, we present it in the following form:

$$\varphi(p) = \int_0^1 \frac{\lambda^{-1}(\lambda-1)^2}{(p + \frac{(1-\lambda)}{\lambda})^2} d\|E_\lambda Sr\|^2. \tag{12.10}$$

Then, its first and second derivatives are written down as follows:

$$\varphi'(p) = -2\int_0^1 \frac{\lambda^{-1}(\lambda-1)^2}{(p + \frac{(1-\lambda)}{\lambda})^3} d\|E_\lambda Sr\|^2, \tag{12.11}$$

$$\varphi''(p) = 6\int_0^1 \frac{\lambda^{-1}(\lambda-1)^2}{(p + \frac{(1-\lambda)}{\lambda})^4} d\|E_\lambda Sr\|^2. \tag{12.12}$$

Owing to the Schwartz inequality

$$\left(\int_0^1 f(\lambda)g(\lambda)d\|E_\lambda Sr\|^2\right)^2 \leq \int_0^1 f^2(\lambda)d\|E_\lambda Sr\|^2 \int_0^1 g^2(\lambda)d\|E_\lambda Sr\|^2$$

and equalities (12.10-12.12) we have

$$\left(\frac{\varphi'(p)}{2}\right)^2 \le \varphi(p)\frac{\varphi''(p)}{6}. \tag{12.13}$$

By definition of the function $\psi(p)$ it follows that $\psi'(p) = -1/2\varphi^{-3/2}\varphi'(p)$, and with the help of (12.10-12.11) we conclude that $\psi(p)$ increases. Then the differentiating we obtain $\psi''(p) = 1/4\psi^{-5/2}(3(\psi'(p))^2 - 2\psi(p)\psi''(p))$ and on the basis of (12.10), (12.13) we deduce that the second derivative is negative, i.e. the function is concave. □

12.3. Methods for Choosing Optimal Parameter

Let us describe two practical methods for choosing the optimal smoothing parameter: the Newton method and 3-d degree Chebyshev method.

12.3.1. Newton Method

Formulas of the Newton method for solving the equation $\psi(p) = \varepsilon^{-1}$ are of the following form:

$$p_{k+1} = p_k - \frac{\psi(p_k) - \varepsilon^{-1}}{\psi'(p_k)}, \tag{12.14}$$

where p_0 is an initial approximation value. The best way is to begin iterations with $p_0 = 0$, the convergence will be provided.

12.3.2. Chebyshev Method of the 3-d Degree

Formulas of the Chebyshev method for solving the equation $\psi(p) = \varepsilon^{-1}$ are of the following form:

$$p_{k+1} = p_k - \frac{\psi(p_k) - \varepsilon^{-1}}{\psi'(p_k)} - \frac{\psi''(p_k)(\psi(p_k) - \varepsilon^{-1})^2}{2(\psi'(p_k))^3}. \tag{12.15}$$

We propose to begin iterations with $p = 0$, although we do not know any substantiation of the convergence of this method for our problem.

12.3.3. Calculating Formulas for Derivatives of $\psi(p)$

Derivatives of the function $\psi(p)$, including those in formulas (12.14-12.15) for the Newton and Chebyshev methods, have to be calculated with the help of the following formulas:

$$\psi(p) = \|A\Sigma_p - z\|_Z^{-1}, \tag{12.16}$$

$$\psi'(p) = -\|A\Sigma_p - z\|_Z^{-3}(A\Sigma_p', A\Sigma_p - z)_Z, \tag{12.17}$$

$$\psi''(p) = 3\|A\Sigma_p - z\|_Z^{-5}(A\Sigma'_p, A\Sigma_p - z)_Z^2 - \\ - \|A\Sigma_p - z\|_Z^{-3}((A\Sigma''_p, A\Sigma_p - z)_Z + (A\Sigma'_p, A\Sigma'_p)_Z), \tag{12.18}$$

where $\Sigma_p = \sigma_{1/p}$. We do not adduce the full proof of these facts, but only give some hints: formula (12.16) clearly follows from (12.4) and (12.7), formulas (12.17-12.18) are obtained by substitution the scalar product instead of the norm in (12.16) and by further differentiation the complicated function.

The formulas used need computation of the derivatives of the spline in the parameter p. In the next three sections we give formulas for the general abstract case and for two particular numerical methods: splines in finite-dimensional subspaces, splines based on reproducing kernels. The main result we come to is in the construction of the formulas for the case $p = 0$. In this case, the smoothing problem reduces to its extremal variant, for which the proposed methods for finding of the spline and its derivatives were unknown before.

12.4. Derivatives of Abstract Smoothing Spline

It is easy to demonstrate that the smoothing spline Σ_p and its derivatives $\Sigma_p^{(k)}$ with $p > 0$ are determined as solutions to the following operator equations:

$$(T^*T + pA^*A)\Sigma_p = pA^*z, \tag{12.19}$$

$$(T^*T + pA^*A)\Sigma'_p = A^*(z - A\Sigma_p), \tag{12.20}$$

$$(T^*T + pA^*A)\Sigma_p^{(k)} = -kA^*A\Sigma_p^{(k-1)}, \quad k \geq 2. \tag{12.21}$$

To realize the algorithm described in Sect. 12.3, we need to determine the spline Σ_0 and its derivatives.

Theorem 12.4. The smoothing splines Σ_p converge with $p \to 0$ to the element $\Sigma_0 \in N(T)$, which is the solution to the following problem

$$\Sigma_0 = \arg \min_{u \in N(T)} \|Au - z\|_Z \tag{12.22}$$

The element Σ_0 may be also determined with the help of the orthogonal conditions

$$(A\Sigma_0 - z, Au)_Z = 0, \quad \forall u \in N(T), \tag{12.23}$$

which (as one may see) are equivalent to problem (12.22).

This Theorem is well known. Note that, according to equation (12.19), the limit spline with $p \to 0$ satisfies the equation $T^*T\Sigma_0 = 0$, i.e. naturally $\Sigma_0 \in N(T)$ (it is known that $N(T) = N(T^*T)$). For further objectives here we formulate and prove a more general Theorem, than the previous one.

Theorem 12.5. Let a sequence of the elements $z_p \in Z$ converge to the element $z \in Z$ with $p \to 0$, i.e. $\|z_p - z\| \to 0$. Then the solutions of problems

$$u_p = \arg\min_{u \in X} \|Tu\|_Y^2 + p\|Au - z_p\|_Z^2. \tag{12.24}$$

or, which is the same, the solutions of operator equations:

$$(T^*T + pA^*A)u_p = pA^*z_p, \tag{12.25}$$

converge with $p \to 0$ to the element $u_0 \in N(T)$, which minimizes the functional $\|Au - z\|_Z$.

Proof. By definition of smoothing splines (12.24) it follows:

$$\|Tu_p\|_Y^2 + p\|Au_p - z_p\|_Z^2 \le p\|Au_0 - z_p\|_Z^2, \quad \forall p \ge 0,$$

because $\|Tu_0\|_Y = 0$. This implies two inequalities

$$\|Tu_p\|_Y^2 \le p\|Au_0 - z_p\|_Z^2, \tag{12.26}$$

$$\|Au_p - z_p\|_Z \le \|Au_0 - z_p\|_Z. \tag{12.27}$$

The sequence $\|u_p\|_X$ is bounded. Naturally, the expression

$$\|Tu_p\|_Y^2 + \|Au_p\|_Z^2$$

is the square of a special norm in X being equivalent to the initial one. This expession is bounded with $0 \le p \le p_0$ owing to (12.26),(12.27) and boundness of the sequence z_p.

In Appendix 1, it is noted that from any bounded sequence one can choose a weakly converging subsequence. Take such a subsequence $\tilde{u}_p$ from u_p. If $\tilde{u}_p \xrightarrow{w} \tilde{u}_0$, then $T\tilde{u}_p \xrightarrow{w} T\tilde{u}_0$, but from (12.26) it follows $\|T\tilde{u}_p\| \to 0$, hence, $T\tilde{u}_0 = 0$. From (12.27) we have:

$$\|A\tilde{u}_p - z_p\| \le \|Au_0 - z\| + \|z_p - z\|. \tag{12.28}$$

The subsequence $A\tilde{u}_p - z_p$ weakly converges to $A\tilde{u}_0 - z$, thus from (12.28) we obtain

$$\|A\tilde{u}_0 - z\| \le \|Au_0 - z\|.$$

Making use of the latter inequality and the equality $T\tilde{u}_0 = 0$, we conclude that $\tilde{u}_0 = u_0$. Consequently, the limit element u_0 does not depend on the subsequence in u_p, hence, the sequence u_p weakly converges to u_0. Since $Au_p - z_p$ weakly converges to $Au_0 - z$ and from (12.28) it follows:

$$\lim_{p \to 0} \|Au_p - z_p\| \le \|Au_0 - z\|,$$

then $Au_p - z_p$ strongly converges to $Au_0 - z$ according to one of the Theorems from Appendix 1. The latter convergence implies the strong convergence of the sequence $A(u_p - u_0)$ to zero. Another sequence $\|Tu_p\|$ also converges to zero, because of (12.26). Finally, from the equivalence of norm it follows

$$c\|u_p - u_0\|_X^2 \le \|Tu_p\|^2 + \|A(u_p - u_0)\|^2 \to 0,$$

i.e. $\|u_p - u_0\|_X \to 0$, when $\alpha \to 0$. □

Theorem 12.6. The smoothing spline Σ_p is a continuously differentiable in the parameter $p \ge 0$. The first derivative Σ_0' satisfies equation

$$T^*T\Sigma_0' = A^*(z - A\Sigma_0), \tag{12.29}$$

more exactly, it is the solution to the following problem:

$$\Sigma_0' = \arg \min_{u \in X,\ T^*Tu = A^*(z - A\Sigma_0)} \|Au\|_Z. \tag{12.30}$$

If u_0 is a particular solution to problem

$$T^*Tu = A^*(z - A\Sigma_0), \tag{12.31}$$

then the first derivative is of the form $\Sigma_0' = u_0 - \pi$, where $\pi \in N(T)$ is determined from equations

$$(A\pi, Au)_Z = (Au_0, Au)_Z, \quad \forall u \in N(T). \tag{12.32}$$

Proof. First, let us explain the sense of (12.30). Any solution to system (12.31) differs from a particular solution u_0 in an element from the kernel of the operator T. The element Σ_0' is such a solution of (12.31), which has the minimum norm after the influence of the operator A. Equations (12.32) are the orthogonal property for variational problem (12.30).

Now let us convince ourselves that equation (12.31) is solvable. It is known that the problem $T^*Tu = f$ is solvable, if $(f, u)_X = 0$ for all $u \in N(T^*T)$. Since $N(T^*T) = N(T)$, then the solvability condition has the form $(A^*(z - A\Sigma_0), u)_X = 0$ for all $u \in N(T)$. It is fulfilled, because it coincides with orthogonal property (12.23). Clearly, this demonstration shows that the element Σ_0', defined in (12.31-12.32) exists and is unique.

Further, we establish that the spline Σ_p is continuously differentiable when $p = 0$ (for $p \ge 0$ this is similarly done). To do this consider, the differences for the derivative $u_p = \frac{\Sigma_p - \Sigma_0}{p}$ and prove that they converge to Σ_0' with $p \to 0$. From equation (12.19) it follows

$$(T^*T + pA^*A)\frac{\Sigma_p - \Sigma_0}{p} = A^*(z - A\Sigma_0), \tag{12.33}$$

and from (12.33) and (12.29) -

$$(T^*T + pA^*A)(u_p - \Sigma_0') = -pA^*A\Sigma_0'.$$

According to Theorem 12.4, the sequence $u_p - \Sigma_0'$ converges to the element from $N(T)$, which minimizes the deviation

$$\|Au + A\Sigma_0'\|^2.$$

From (12.30) it follows that this element must be zero. Thus, the divided differences of u_p converge to Σ_0'. Additionally, we have to convince ourselves that the derivative Σ_0' is continuous for $p = 0$. For this purpose consider the equation

$$(T^*T + pA^*A)(\Sigma_p' - \Sigma_0') = -pA^*A\left(\Sigma_0' + \frac{\Sigma_p - \Sigma_0}{p}\right),$$

which follows from (12.20) and (12.29). The elements $z_p = -A\left(\Sigma_0' + \frac{\Sigma_p - \Sigma_0}{p}\right)$ converge to $-2A\Sigma_0$, hence, owing to Theorem 12.5 the difference $u_p = \Sigma_p' - \Sigma_0'$ converges to the element $u \in N(T)$, which minimizes the deviation $\|Au + 2A\Sigma_0'\|^2$. From (12.30) it follows, that such an element is zero. The Theorem is proved. □

Theorem 12.7. The smoothing spline Σ_p is infinitely differentiable on the parameter $p \geq 0$. The derivatives of the spline $\Sigma_0^{(k)}$ for $k \geq 2$ satisfy the equality

$$T^*T\Sigma_0^{(k)} = -kA^*A\Sigma_0^{(k-1)}, \tag{12.34}$$

more exactly, they are the solutions to the following problems:

$$\Sigma_0^{(k)} = \arg \min_{u \in X,\ T^*Tu = -kA^*A\Sigma_0^{(k-1)}} \|Au\|_Z. \tag{12.35}$$

If u_0 is a particular solution to the problem

$$T^*Tu = -kA^*A\Sigma_0^{(k-1)}, \tag{12.36}$$

then the respective derivative is of the form $\Sigma_0^{(k)} = u_0 - \pi$, where $\pi \in N(T)$ is determined from the equations

$$(A\pi, Au)_Z = (Au_0, Au)_Z, \quad \forall u \in N(T). \tag{12.37}$$

Proof. We will follow the same way as in the previous theorem, but omit the most of the explanations. Firstly, let us convince ourselves that equation (12.36) is solvable. The solvability condition is of the form: $(A^*A\Sigma_0^{(k-1)}, u)_X = 0$ for all $u \in N(T)$. For $k = 2$ it is fulfilled in consequence of orthogonal properties (12.32), and for $k \geq 2$ - in consequence of (12.37), respectively.

Now we establish continuous differentiability only for $p = 0$. Consider the differences for the k-th derivative:

$$u_p = \frac{\Sigma_p^{(k-1)} - \Sigma_0^{(k-1)}}{p}$$

and prove that they converge to $\Sigma_0^{(k)}$ with $p \to 0$. From equations (12.20-12.21), (12.29), (12.34) it follows that

$$(T^*T + pA^*A)u_p = -kA^*A\Sigma_0^{(k-1)}$$

and

$$(T^*T + pA^*A)(u_p - \Sigma_0^{(k)}) = -pA^*A\Sigma_0^{(k)}.$$

By Theorem 12.4 the sequence $u_p - \Sigma_0^{(k)}$ converges to an element from $N(T)$, which minimizes the deviation $\|Au + A\Sigma_0^{(k)}\|^2$. Property (12.3) implies that this element has to be zero. Thus, the divided differences u_p converge to $\Sigma_0^{(k)}$. Yet we must check that the derivative $\Sigma_p^{(k)}$ is continuous for $p = 0,\ k \geq 2$. To do this, consider the equation

$$(T^*T + pA^*A)(\Sigma_p^{(k)} - \Sigma_0^{(k)}) = -pA^*A\left(\Sigma_0^{(k)} + k\frac{\Sigma_p^{(k-1)} - \Sigma_0^{(k-1)}}{p}\right),$$

which follows from (12.21), (12.34). The elements $z_p = A(\Sigma_0' + k\frac{\Sigma_p - \Sigma_0}{p})$ converge to $-(k+1)A\Sigma_0'$, consequently, by Theorem 12.5 the difference $u_p = \Sigma_p^{(k)} - \Sigma_0^{(k)}$ converges to an element $u \in N(T)$, which minimizes the deviation

$$\|Au + (k+1)A\Sigma_0'\|^2.$$

From (12.35) it follows that such an element has to be zero. □

12.5. Derivaties of the Smoothing Spline on Subspace

Let E be a finite-dimensional subspace in X having the basis $\omega_1, ..., \omega_N$. Note (see Chapter 4), that we call the element

$$\sigma_\alpha = \sum_{i=1}^{n} \sigma_{i,\alpha}\omega_i \tag{12.38}$$

as a smoothing spline on subspace, if it minimizes the functional $\alpha\|Tu\|_Y^2 + \|Au - z\|_Z^2$ among the elements from E. The vector of the coefficients $\bar\sigma_\alpha = (\sigma_{1,\alpha}, ..., \sigma_{n,\alpha})^T$ is determined from the system of linear algebraic equations (SLAE)

$$(\alpha\bar T + \bar A)\bar\sigma_\alpha = f, \tag{12.39}$$

where $\bar T$ is a matrix with the common element $t_{ij} = (T\omega_i, T\omega_j)_Y$, $\bar A$ is the matrix $a_{ij} = (A\omega_i, A\omega_j)_Z$, f is the vector $(A\omega_i, z)$. Introduce the notations $\Sigma_p = \sigma_{1/p}$, $\bar\Sigma_p = \bar\sigma_{1/p}$. From (12.39) it is clear how to obtain SLAE for the derivatives of the spline on subspace with $p > 0$:

$$(\bar{T} + p\bar{A})\bar{\Sigma}_p = pf, \tag{12.40}$$
$$(\bar{T} + p\bar{A})\bar{\Sigma}_p' = f - \bar{A}\bar{\Sigma}_p, \tag{12.41}$$
$$(\bar{T} + p\bar{A})\Sigma_p^{(k)} = -k\bar{A}\bar{\Sigma}_p^{(k-1)}, \quad k \geq 2. \tag{12.42}$$

To find the spline and its derivatives for $p = 0$, assume that $N(T)$ is contained in E and has the basis $e_1, ..., e_m$.

Theorem 12.8. The following three statements are valid:

1. The spline Σ_0 is of the form

$$\Sigma_0 = \sum_{i=1}^{m} q_i e_i, \tag{12.43}$$

whose coefficients $q_1, ..., q_m$ are determined from the following SLAE:

$$\sum_{i=1}^{m} (Ae_i, Ae_j)_Z \cdot q_i = (z, Ae_j)_Z, \quad j = 1, ..., m.$$

2. Let $u = (u_1, ..., u_n)^T$ be a solution to the following problem:

$$\bar{T}u = f - \bar{A}\bar{\Sigma}_0, \tag{12.44}$$

where $\bar{\Sigma}_0$ is a decomposition of the spline Σ_0 on the basis $\omega_1, ..., \omega_n$. The first derivative of the smoothing spline Σ_0' is of the form

$$\Sigma_0' = \sum_{i=1}^{n} u_i \omega_i - \sum_{i=1}^{m} q_i e_i, \tag{12.45}$$

whose coefficients $q_1, ..., q_m$ are determined from SLAE:

$$\sum_{i=1}^{m} (Ae_i, Ae_j)_Z \cdot q_i = (Au, Ae_j)_Z, \quad j = 1, ..., m. \tag{12.46}$$

3. Let $u = (u_1, ..., u_n)^T$ be a particular solution to the problem:

$$\bar{T}u = -k\bar{A}\bar{\Sigma}_0^{(k-1)}, \tag{12.47}$$

where $\bar{\Sigma}_0^{(k-1)}$ is the decomposition of the spline $\Sigma_0^{(k-1)}$ on the basis $\omega_1, ..., \omega_n$ for $k \geq 2$. Then, k-th derivative of the smoothing spline Σ_0^k is of the form

$$\Sigma_0^k = \sum_{i=1}^{n} u_i \omega_i - \sum_{i=1}^{m} q_i e_i, \tag{12.48}$$

whose coefficients $q_1, ..., q_m$ are determined from SLAE:

$$\sum_{i=1}^{m} (Ae_i, Ae_j)_Z \cdot q_i = (Au, Ae_j)_Z, \quad j = 1, ..., m. \tag{12.49}$$

Proof. In this theorem, the first statement follows from Theorem 12.4, the second and third statements follow from Theorem 12.6 and 12.7, respectively. □

12.6. Derivatives of the Smoothing Spline by Reproducing Kernels

Recall (see Chapter 2), that these splines are determined from the following SLAEs

$$\begin{bmatrix} \alpha I + G & B \\ B^* & 0 \end{bmatrix} \begin{bmatrix} \lambda_\alpha \\ \nu_\alpha \end{bmatrix} = \begin{bmatrix} z \\ 0 \end{bmatrix},$$

where I is the identity matrix. Setting $\Lambda_p = \lambda_{1/p}$, $M_p = \nu_{1/p}$ one can assure oneself, that the spline $\Sigma_p = (\Lambda_p, M_p)$ and its derivatives are determined from the following SLAEs

$$\begin{bmatrix} I + pG & pB \\ pB^* & 0 \end{bmatrix} \begin{bmatrix} \Lambda_p \\ M_p \end{bmatrix} = \begin{bmatrix} pz \\ 0 \end{bmatrix}, \tag{12.50}$$

$$\begin{bmatrix} I + pG & pB \\ pB^* & 0 \end{bmatrix} \begin{bmatrix} \Lambda'_p \\ M'_p \end{bmatrix} = \begin{bmatrix} z \\ 0 \end{bmatrix} - \begin{bmatrix} G & B \\ B^* & 0 \end{bmatrix} \begin{bmatrix} \Lambda_p \\ M_p \end{bmatrix}, \tag{12.51}$$

$$\begin{bmatrix} I + pG & pB \\ pB^* & 0 \end{bmatrix} \begin{bmatrix} \Lambda_p^{(k)} \\ M_p^{(k)} \end{bmatrix} = -k \begin{bmatrix} G & B \\ B^* & 0 \end{bmatrix} \begin{bmatrix} \Lambda_p^{(k-1)} \\ M_p^{(k-1)} \end{bmatrix}, \quad k \geq 2. \tag{12.52}$$

Theorem 12.9. The smoothing spline Σ_0 and its derivatives for $p = 0$ are determined from the following SLAE

$$\Lambda_0 = 0, \quad B^*BM_0 = B^*z, \tag{12.53}$$

$$\Lambda'_0 = z - BM_0, \quad B^*BM'_0 = -B^*G\Lambda'_0, \tag{12.54}$$

$$\Lambda_0^{(k)} = -k(G\Lambda_0^{(k-1)} + BM_0^{(k-1)}), \quad B^*BM_0^{(k)} = -B^*G\Lambda_0^{(k)}. \tag{12.55}$$

Proof. Making use of the limit with $p \to 0$ in equations (12.50-12.52) one can make certain, that the equalities for Λ_0, Λ'_0 and $\Lambda^{(k)}$ with $k \geq 2$ in (12.53-12.55) are valid. The equalities for M_0, M'_0 and $M_0^{(k)}$ with $k \geq 2$ in (12.53-12.55) follow from conditions (12.23), (12.32) and (12.37). □

12.7. Numerical Formulas for Optimal Smoothing Parameter for Different Algorithms

12.7.1. Spline on Subspaces

When solving the equation $\psi(p) = \varepsilon^{-1}$ by the Newton method

$$p_{k+1} = p_k - \frac{\psi(p_k) - \varepsilon^{-1}}{\psi'(p_k)}, \quad k \in 0, ..., N \tag{12.56}$$

or by the 3-d degree Chebyshev method

$$p_{k+1} = p_k - \frac{\psi(p_k) - \varepsilon^{-1}}{\psi'(p_k)} - \frac{\psi''(p_k)(\psi(p_k) - \varepsilon^{-1})^2}{2(\psi'(p_k))^3}, \tag{12.57}$$

set $p = 0$.

Derivatives of the function $\psi(p)$ presented in formulas (12.56-12.57), should be calculated by the following formulas:

$$\psi(p) = \|A\Sigma_p - z\|_Z^{-1}, \tag{12.58}$$

$$\psi'(p) = -\|A\Sigma_p - z\|_Z^{-3}(\bar{\Sigma}'_p, \bar{A}\bar{\Sigma}_p - f), \tag{12.59}$$

$$\begin{aligned} \psi''(p) &= 3\|A\Sigma_p - z\|_Z^{-5}(\bar{\Sigma}'_p, \bar{A}\bar{\Sigma}_p - f)^2 - \\ &- \|A\Sigma_p - z\|_Z^{-3}((\bar{A}\bar{\Sigma}''_p, \bar{\Sigma}_p - f) + (\bar{A}\bar{\Sigma}'_p, \bar{\Sigma}'_p)), \end{aligned} \tag{12.60}$$

where

$$\|A\Sigma_p - z\|_Z = ((\bar{A}\bar{\Sigma}_p, \bar{\Sigma}_p) - 2(\bar{\Sigma}_p, f) + \|z\|_Z^2)^{1/2}.$$

Here we have used the notation from Sect. 12.5. An the first step with $p = 0$ the vector of coefficients $\bar{\Sigma}_0$ and its derivatives should be calculated with the help of the algorithm, proposed in Theorem 12.8 and then one should use formulas (12.40-12.42).

12.7.2. Splines on the Basis of Reproducing Kernels

Solving the equation $\psi(p) = \varepsilon^{-1}$ by the Newton method (12.56) or by the Chebyshev method (12.57), one should calculate the derivatives of the function $\psi(p)$ by the following formulas

$$\begin{aligned} \psi(p) &= \|G\Lambda_p + BM_p - z\|_Z^{-1}, \\ \psi'(p) &= -\|G\Lambda_p + BM_p - z\|_Z^{-3}(G\Lambda'_p + BM'_p, G\Lambda_p + BM_p - z)_Z, \\ \psi''(p) &= 3\|G\Lambda_p + BM_p - z\|_Z^{-5}(G\Lambda'_p + BM'_p, G\Lambda_p + BM_p - z)_Z^2 - \\ &- \|G\Lambda_p + BM_p - z\|_Z^{-3}((G\Lambda''_p + BM''_p, G\Lambda_p + BM_p - z)_Z + \\ &+ (G\Lambda'_p + BM'_p, G\Lambda'_p + BM'_p)_Z). \end{aligned}$$

Here we have used the notations of Sect.12.6. At the first step with $p = 0$ the vector of coefficients (Λ_0, M_0) and its derivatives should be calculated with the help of formulas (12.53-12.55), and then, for $p > 0$ - by formulas (12.50-12.52).

12.8. Uniform Convergence of the Taylor Series for Smoothing Splines

12.8.1. Investigation of Taylor Series by Parameter α

When computing the optimal smoothing parameter one needs to know how to calculate the spline and its derivatives for different smoothing parameters. In this Section, we fix the parameter $\alpha_0 > 0$ and using some auxiliary assertions reduce the problem of calculating the spline and its derivatives for different smoothing parameters to a few smoothing problems with the same parameter α_0. In order to do it, investigate the convergence of the Taylor series. It allows us to approximately find the splines from the same operator equation with different right-hand sides.

The derivatives $\sigma_{\alpha_0}^{(k)}$ exist for all $k \geq 0$, thus there is a point to consider the partial sums of the Taylor series

$$S_{n,a} = \sum_{k=0}^{n} \frac{(\alpha - \alpha_0)^k}{k!} \sigma_{\alpha_0}^{(k)}. \tag{12.61}$$

Lemma 12.1. The partial sums of the Taylor series $S_{n,\alpha}$ and the remaining terms

$$R_{n,\alpha} = \sigma_\alpha - S_{n,\alpha} \tag{12.62}$$

belong to the space of splines $Sp(T, A)$.

Proof. With the help of induction, it is not difficult to verify that for any smoothing parameter and for any derivative $k \geq 1$, the following equalities

$$(\alpha T^* T + A^* A)\sigma_\alpha^{(k)} = A^* z_k,$$

take place, if

$$z_k = k \frac{A\sigma_\alpha^{(k-1)} - z_{k-1}}{\alpha} + z'_{k-1},$$

$k = 1, 2, \ldots;\ \sigma_\alpha^{(0)} = \sigma_\alpha,\ z_0 = z$. Thus, all derivatives of the smoothing spline $\sigma_{\alpha_0}^{(k)}$ lie in the space of splines, consequently, their linear combinations $S_{n,\alpha}$ also lie in $Sp(T, A)$. The element σ_α is a spline, consequently, $R_{n,\alpha}$ is a spline as a difference of two splines. □

Lemma 12.2. The following recurrent formula

$$(\alpha_0 T^* T + A^* A) S_{n,\alpha} = (\alpha_0 - \alpha) T^* T S_{n-1,\alpha} + A^* z, \tag{12.63}$$

with the initial data $S_{0,\alpha} = \sigma_\alpha$ may be used for calculating the partial sums of the Taylor series for a smoothing spline.

Proof. Making use of the operator equations for spline

$$(\alpha T^*T + A^*A)\sigma_\alpha = A^*z, \tag{12.64}$$

and for its derivatives

$$(\alpha T^*T + A^*A)\sigma_\alpha^{(k)} = -kT^*T\sigma_\alpha^{(k-1)},$$

and, also, utilizing formula (12.61), we have a few equalities:

$$(\alpha_0 T^*T + A^*A)S_{n,\alpha} = (\alpha_0 T^*T + A^*A)\sum_{k=0}^{n} \frac{(\alpha-\alpha_0)^k}{k!}\sigma_{\alpha_0}^{(k)} =$$

$$= \sum_{k=1}^{n} \frac{(\alpha-\alpha_0)^k}{k!}(-kT^*T\sigma_{\alpha_0}^{(k-1)}) + A^*z = (\alpha_0 - \alpha)T^*TS_{n-1,\alpha} + A^*z,$$

which finish the proof of the lemma. □

Lemma 12.3. Reccurent formula (12.63) and the following recurrent formula

$$(\alpha_0 T^*T + A^*A)S'_{n,\alpha} = (\alpha_0 - \alpha)T^*TS'_{n-1,\alpha} - T^*TS_{n-1,\alpha}, \tag{12.65}$$

with the initial data $S_{0,\alpha} = \sigma_{\alpha_0}$, $S'_{0,\alpha} = \sigma'_{\alpha_0}$ may be used for calculating the derivatives of the partial sums of the Taylor series.

Proof. This formula is established by differentiating formula (12.63). □

Theorem 12.10. Partial sums of Taylor series (12.61) uniformly converge to the spline σ_α for all closed subintervals of the interval $\alpha \in (0, 2\alpha_0)$, i.e.

$$\sigma_\alpha = \sum_{k=0}^{\infty} \frac{(\alpha-\alpha_0)^k}{k!}\sigma_{\alpha_0}^{(k)}. \tag{12.66}$$

If the unit ball of the space $Sp(T, A)$ is compact in X, then the uniform convergence takes place for the whole closed interval $\alpha \in [0, 2\alpha_0]$.

Proof. From equality (12.64) follows

$$(\alpha_0 T^*T + A^*A)\sigma_\alpha = (\alpha_0 - \alpha)T^*T\sigma_\alpha + A^*z,$$

and from the latter equality and from (12.63) follows

$$(\alpha_0 T^*T + A^*A)(\sigma_\alpha - S_{n,\alpha}) = (\alpha_0 - \alpha)T^*T(\sigma_\alpha - S_{n-1,\alpha}),$$

or, taking into account the notation of remaining terms, -

$$R_{n,\alpha} = [(\alpha_0 T^*T + A^*A)^{-1}(\alpha_0 - \alpha)T^*T]R_{n-1,\alpha}. \tag{12.67}$$

In order that the Taylor series be able to converge, it is necessary that the spectral radii $\rho(M)$ of the operator $M = (\alpha_0 T^*T + A^*A)^{-1}(\alpha_0 - \alpha)T^*T$ should not exceed one. For estimating the spectral radii, the following inequality

$$\rho(M) \leq \sup_{\|u\|=1} \frac{|(DMu, u)_X|}{|(Du, u)_X|},$$

can be chosen with any positive operator $D : X \to X$. Let $D = (\alpha_0 T^*T + A^*A)$. Then we have

$$\rho(M) = \sup_{\|u\|=1} \frac{|(\alpha_0 - \alpha)T^*Tu, u)_X|}{|((\alpha_0 T^*T + A^*A)u, u)_X|} = \\ = \frac{|\alpha_0 - \alpha|}{\alpha_0} \sup_{\|u\|=1} \frac{\alpha_0\|Tu\|_Y^2}{\alpha_0\|Tu\|_Y^2 + \|Au\|_Z^2}. \tag{12.68}$$

Clearly, the supremum does not exceed one. Thus,

$$\rho(M) \leq \frac{|\alpha_0 - \alpha|}{\alpha_0},$$

and if the smoothing parameter α changes its value in a closed subinterval of $(0, 2\alpha_0)$, then the spectral radias is bounded by unity, which provides the uniform convergence in this subinterval.

In order to prove the convergence of the series for $\alpha = 0$ and $\alpha = 2\alpha_0$, if the additional condition takes place, note that in recurrent formulas (12.67) the iterative elements (remaining terms) belong to the space of splines. Use the notation $M|_{Sp(T,A)}$ instead of the restriction of the operator M on the space $Sp(T, A)$. If we prove the weaker condition

$$\rho(M|_{Sp(T,A)}) < 1,$$

then it will provide for the convergence of the Taylor series (12.61). Supremum in (12.68) is exactly less than unity, if it is considered on a compact of $Sp(T, A)$. In this case, the spectral radius is uniformely restricted by one on the total interval $[0, 2\alpha_0]$, and, even, on a little greater interval, and the theorem is proved. □

Theorem 12.11. If for $\alpha = 0$, series (12.66) is converging, then it converges to the interpolation spline.

Proof. Naturally, since the space $Sp(T, A)$ is closed, then the limit of partial sums, which is the sum of the series, must be the element of $Sp(T, A)$. Making recurrent formula (12.63) tend to its limit, one can obtain the equality $A^*A\sigma_0 = A^*z$, which implies $A\sigma_0 = z$, i.e. the spline σ_0 satisfies the interpolating conditions. □

12.8.2. Investigation of Taylor Series on Parameter p

A very interesting question arises about convergence of the Taylor series for the smoothing spline Σ_p at the point $p = 0$. Does a nonzero radius of convergence of this series exist? The reply to this question is given in the following theorem.

Theorem 12.12. The partial sums:

$$S_{n,p} = \sum_{k=0}^{n} \frac{p^k}{k!} \Sigma_0^{(k)}$$

uniformly converge to the splines Σ_p on some closed interval $[0, p_0]$, where $p_0 > 0$.

Proof. Making use of Theorems from Section 12.4 about the representation of the abstract smoothing spline Σ_p with $p = 0$, we have

$$\begin{aligned} T^*TS_{n,p} &= T^*T\Sigma_0 + pT^*T\Sigma_0' + T^*T\sum_{k=2}^{n} \frac{p^k}{k!}\Sigma_0^{(k)} = \\ &= pA^*(z - A\Sigma_0) - p\sum_{k=2}^{n} \frac{p^{k-1}}{(k-1)!} A^*A\Sigma_0^{(k-1)} = \\ &= pA^*(z - AS_{n-1,p}). \end{aligned}$$

Thus, we obtain the following recurrent formula for the calculation of the partial sums

$$T^*TS_{n,p} = pA^*(z - AS_{n-1,p}). \tag{12.70}$$

Taking into account the equation for the spline Σ_p, we obtain

$$(T^*T + pA^*A)(S_{n,p} - \Sigma_p) = pA^*A(S_{n,p} - S_{n-1,p}).$$

From here it follows that the difference $S_{n,p} - \Sigma_p$ converges to zero, if the common element of the Taylor series:

$$\frac{p^k}{k!}\Sigma_0^{(k)} = S_{k,p} - S_{k-1,p}$$

converges to zero. From equalities (12.29), (12.34) we have

$$\begin{aligned} &\frac{p^k}{k!}\Sigma_0^{(k)} = (T^*T)^{-1}A^*A\frac{p^k}{(k-1)!}\Sigma_0^{(k-1)} = \dots = \\ &(p(T^*T)^{-1}A^*A)^{k-1}p\Sigma_0'. \end{aligned} \tag{12.71}$$

Here $(T^*T)^{-1}$ stands for the inverse operator to (T^*T) on the subspace of the element u_0, satisfying property (12.37). This subspace is orthogonal to $N(T)$ in a special scalar product, thus, the inverse operator $(T^*T)^{-1}$ exists and is bounded. Clearly, common element (12.71) converges to zero, if

$$p < \rho((T^*T)^{-1}A^*A)^{-1},$$

where ρ is the function of the spectral radius of operators. □

12.9. Discussion of Benefits of Extrapolation for Spline Construction on Convex Set

A particular case of the spline problem on a convex set is the interpolating spline problem. Using the method of splines in subspaces increases SLAE up to two times exactly, which negatively affects the rate and accuracy of spline construction. Besides, SLAE ceases to be positively defined. Using extrapolation of two or more smoothing splines with different smoothing parameters

$$\sigma_0 \cong a_1\sigma_{\alpha_1} + a_2\sigma_{\alpha_2} + a_3\sigma_{\alpha_3}$$

allows us to avoid these difficulties. But, it is necessary to solve a few smoothing problems. With the help of the idea of extrapolation on the basis of the Taylor series, proposed in this Chapter, the requirement may be removed in a sense. In conformity with formula (12.66) the interpolating spline is represented (under certain restrictments, which are in general fulfilled) in the following form

$$\sigma_0 = \sum_{k=0}^{\infty}(-1)^k \frac{\alpha_0^{(k)}}{k!}\sigma_{\alpha_0}^{(k)}.$$

Consequently, by Lemma 12.2 (Sect. 12.8) the interpolating spline may be found as the limit of the sequence of the partial sums $S_{n,0}$, defined from the following smoothing problems with the fixed smoothing parameter α_0:

$$(\alpha_0 T^*T + A^*A)S_{n,0} = \alpha_0 T^*TS_{n-1,0} + A^*z,$$

where $S_{0,0} = \sigma_{\alpha_0}$. Here $S_{n,0}$ is n-th partial sum of the Taylor series. Now one make of using the ideas of the full and incomplete factorization in order to take advantages from solving a few SLAE with a fixed matrix.

The same idea may be used, when one finds an arbitrary spline on a convex set. In accordance with the algorithm proposed in Sect. 12.7, beginning with the second iteration of the Newton method or the Chebyshev method, there arises a sequence of smoothing problems with the decreasing parameters $\alpha_1 = 1/p_1, \alpha_2 = 1/p_2$, and so on. If here we fix the smoothing parameter $\alpha = \alpha_1$, then the rest problems with different smoothing parameters may be found as limits of solutions to the problems with this fixed parameter. Since the iterative smoothing parameters will be less than α_1, then the convergence takes place, and the respective limit exists. Thus, there arises an idea to use in the formulas of the Newton methods (or the Chebyshev method of the third degree) instead of the exact spline and its derivatives the approximate ones, calculated by formulas (12.63), (12.65). This, again, allows one to get advantages from the solution to few systems with a fixed matrix.