Chapter 10

Iterative methods

Contents

10.1
Introduction

In introductory classes, one occasionally maximizes or minimizes a function by taking the derivative, setting it to zero, and solving. In n-dimensional problems this generalizes to solving the $n \times n$ system of equations

$$\nabla \Phi(\boldsymbol{x}) = \mathbf{0}.$$

In many real-world problems, this system of equations has no analytical solution, so numerical methods are required. Usually such methods are **iterative**: we start with an initial guess $\boldsymbol{x}_0$ of the solution, from that generate a new guess $\boldsymbol{x}_1$, and so on. A good iterative algorithm will rapidly converge to a solution of the system of equations.

Even with "only" the material in chapter 2 as background, we can already study some types of iterative methods in considerable generality, *i.e.*, covering both finite-dimensional *and* infinite dimensional cases. All we need is a normed space (or Banach space).

What we cannot discuss (rigorously) yet the large family of iterative methods that uses **derivatives** of a cost function, since we will not define derivatives until Ch. 7.

Subtracting each side of the above equality yields the equivalent condition

$$\boldsymbol{x} = \boldsymbol{x} - \nabla \Phi(\boldsymbol{x})$$

or equivalently

$$\boldsymbol{x} = T(\boldsymbol{x}),$$

where $T(\boldsymbol{x}) = \boldsymbol{x} - \nabla \Phi(\boldsymbol{x})$.

At this point, we disregard Φ, and focus on the problem of solving the system of equations $\boldsymbol{x} = T(\boldsymbol{x})$. Such problems arise in many applications.

Later we will return to the minimization problem, after we have defined derivatives in the context of general normed spaces.

10.2
Successive approximation

Definition. A solution to the system of equations $\boldsymbol{x} = T(\boldsymbol{x})$ is called a **fixed point** of T, since such an $\boldsymbol{x}$ is "fixed" under T.

The recursive form of the equality $\boldsymbol{x} = T(\boldsymbol{x})$ very strongly suggests the following iterative algorithm:

$$\boldsymbol{x}_{n+1} = T(\boldsymbol{x}_n),$$

which is called the **method of successive approximation** (or substitution).

The key questions to investigate are the following.

- Does a fixed point of T exist?
- If so, is it unique?
- Does $\{\boldsymbol{x}_n\}$ converge to a fixed point of T?

The answers to the above questions depend both on T, of course, as well as the norm of the underlying normed space, since convergence is always defined with respect to some norm.

Contraction mappings

Definition. If S is a subset of a normed space $(\mathcal{X}, \|\cdot\|)$, and T is a transformation from S into S, then we call T a **contraction mapping** on S iff $\exists \alpha \in [0,1)$ such that

$$\|T(\boldsymbol{x}) - T(\boldsymbol{y})\| \leq \alpha \|\boldsymbol{x} - \boldsymbol{y}\|, \ \forall \boldsymbol{x}, \boldsymbol{y} \in S.$$

Later we will see that if T is appropriately differentiable then it suffices to have $\|T'(\boldsymbol{x})\| \leq \alpha < 1$.

Fact. If T is a contraction mapping, then T is **continuous**[1].

Example. $T(\boldsymbol{x}) = \boldsymbol{x} + \frac{1}{2}(\boldsymbol{z} - \boldsymbol{x})$ is a contraction mapping.

Example. $T(\boldsymbol{x}) = \alpha \boldsymbol{x}$ is not a contraction mapping for $\alpha \geq 1$.

Theorem. *(Contraction mapping theorem)*
Let T be a contraction mapping on a complete subset S of a normed space $(\mathcal{X}, \|\cdot\|)$.
- *There is a unique fixed point $\boldsymbol{x}_\star \in S$ satisfying $\boldsymbol{x}_\star = T(\boldsymbol{x}_\star)$.*
- *The method of successive approximations yields a sequence $\{\boldsymbol{x}_n\}$ that converges to $\boldsymbol{x}_\star$, for any initial guess $\boldsymbol{x}_0 \in S$.*

Proof.

$$\|\boldsymbol{x}_{n+1} - \boldsymbol{x}_n\| = \|T(\boldsymbol{x}_n) - T(\boldsymbol{x}_{n-1})\| \leq \alpha \|\boldsymbol{x}_n - \boldsymbol{x}_{n-1}\|,$$

so by induction

$$\|\boldsymbol{x}_{n+1} - \boldsymbol{x}_n\| \leq \alpha^n \|\boldsymbol{x}_1 - \boldsymbol{x}_0\|.$$

(This is not quite Cauchy, but it is getting close. So now we try to show that $\{\boldsymbol{x}_n\}$ is Cauchy.)

. . . .

Claim 1. $\{\boldsymbol{x}_n\}$ is Cauchy.
(This is the natural thing to try in a complete subset since it is usually easier than showing convergence directly.)

Consider $\|\boldsymbol{x}_m - \boldsymbol{x}_n\| = \|\boldsymbol{x}_{n+p} - \boldsymbol{x}_n\|$ where $m = n + p$ and $p > 0$ w.l.o.g. Then

$$\begin{aligned}
\|\boldsymbol{x}_m - \boldsymbol{x}_n\| &= \|\boldsymbol{x}_{n+p} - \boldsymbol{x}_n\| \leq \|\boldsymbol{x}_{n+p} - \boldsymbol{x}_{n+p-1}\| + \cdots + \|\boldsymbol{x}_{n+1} - \boldsymbol{x}_n\| \\
&\leq \left(\alpha^{n+p-1} + \cdots + \alpha^n\right) \|\boldsymbol{x}_1 - \boldsymbol{x}_0\| \leq \alpha^n \sum_{k=0}^{\infty} \alpha^k \|\boldsymbol{x}_1 - \boldsymbol{x}_0\| = \frac{\alpha^n}{1-\alpha} \|\boldsymbol{x}_1 - \boldsymbol{x}_0\|.
\end{aligned}$$

This can be made arbitrarily small for n sufficiently large, so $\{\boldsymbol{x}_n\}$ is Cauchy.

Since S is complete, the Cauchy sequence $\{\boldsymbol{x}_n\}$ converges to some limit $\boldsymbol{x}_\star \in S$.

. . . .

Claim 2. $\boldsymbol{x}_\star = T(\boldsymbol{x}_\star)$

Since T is a contraction mapping and hence continuous:

$$\boldsymbol{x}_\star = \lim_{n\to\infty} \boldsymbol{x}_n = \lim_{n\to\infty} T(\boldsymbol{x}_{n-1}) = T\left(\lim_{n\to\infty} \boldsymbol{x}_{n-1}\right) = T(\boldsymbol{x}_\star).$$

. . . .

Claim 3. $\boldsymbol{x}_\star$ is unique.

Suppose $\boldsymbol{y} = T(\boldsymbol{y})$ also. Then

$$\|\boldsymbol{x}_\star - \boldsymbol{y}\| = \|T(\boldsymbol{x}_\star) - T(\boldsymbol{y})\| \leq \alpha \|\boldsymbol{x}_\star - \boldsymbol{y}\|,$$

but $\alpha < 1$, so $\|\boldsymbol{x}_\star - \boldsymbol{y}\| = 0$ hence $\boldsymbol{x}_\star = \boldsymbol{y}$. □

. . . .

Remark. The book's theorem statement says "$\boldsymbol{x}_\star$ can be obtained" by successive approximations. It is not defined what it means to "obtain" $\boldsymbol{x}_\star$. In practice we definitely never really "obtain" $\boldsymbol{x}_\star$, we just run the algorithm several iterations and hope that the "final" $\boldsymbol{x}_n$ is "close enough" to $\boldsymbol{x}_\star$ for practical purposes.

See text for illustrations of successive approximations in 1D.

---

[1] In fact, such a T is "more than" continuous, it is a (globally) **Lipschitz continuous** mapping [3, p. 68], which is a stronger condition than mere continuity.

There are many generalizations of the method of successive approximations. Here is one of them. (Used in the diffeq application!)

Theorem. *Suppose S is a complete subset in $(\mathcal{X}, \|\cdot\|)$, and $T : S \to S$ is continuous, and T^m is a contraction mapping on S for some $m \in \mathbb{N}$.*
- *There is a unique fixed point $\boldsymbol{x}_\star \in S$ satisfying $\boldsymbol{x}_\star = T(\boldsymbol{x}_\star)$.*
- *The method of successive approximations yields a sequence $\{\boldsymbol{x}_n\}$ that converges to $\boldsymbol{x}_\star$, for any initial guess $\boldsymbol{x}_0 \in S$.*

Proof. Consider the subsequence $\{\boldsymbol{x}_{mk+j}\}_{k=1}^\infty$ for $j = 1, 2, \ldots, m$.

Claim 1. $\{\boldsymbol{x}_{mk+j}\}_{k=1}^\infty$ converges to a point $\bar{\boldsymbol{x}}_j$ in S.

Observe $\boldsymbol{x}_{m(k+1)+j} = T^m(\boldsymbol{x}_{mk+j})$, where T^m is a contraction and S is complete. So by the contraction mapping theorem, $\{\boldsymbol{x}_{mk+j}\}_{k=1}^\infty$ converges to a unique $\bar{\boldsymbol{x}}_j \in S$.

Claim 2. $\bar{\boldsymbol{x}}_1 = \cdots = \bar{\boldsymbol{x}}_m$

Each $\bar{\boldsymbol{x}}_j$ is a fixed point of T^m, but since T^m is a contraction mapping it has a unique fixed point by the contraction mapping theorem. Let $\boldsymbol{x}_\star = \bar{\boldsymbol{x}}_1$.

Claim 3. $\boldsymbol{x}_n \to \boldsymbol{x}_\star$.

For all $\varepsilon > 0$, and $j = 1, \ldots, m$, $\exists K_j$ s.t. $k > K_j \Longrightarrow \|\boldsymbol{x}_{mk+j} - \boldsymbol{x}_\star\| < \varepsilon$.
Let $K = \max\{K_1, \ldots, K_m\}$ and $N = mK + 1$, then $n > N \Longrightarrow \|\boldsymbol{x}_n - \boldsymbol{x}_\star\| < \varepsilon$. Since ε was arbitrary, $\boldsymbol{x}_n \to \boldsymbol{x}_\star$.

Claim 4. $\boldsymbol{x}_\star = T(\boldsymbol{x}_\star)$.
Since T is continuous, see Claim 2 of CMT.

Claim 5. $\boldsymbol{x}_\star$ is the unique fixed point of T in S.

Suppose $\boldsymbol{y} = T(\boldsymbol{y})$. Then (recursing) $\boldsymbol{y} = T^m(\boldsymbol{y})$, so $\boldsymbol{y}$ is a fixed point of T^m.
But T^m is a contraction with a unique fixed point, so $\boldsymbol{y} = \boldsymbol{x}_\star$. □

....

(This style of proof, which I have borrowed from Prof. Grizzle's notes, where each sub-claim is clearly identified and proven, is a little less succinct as Luenberger's, but very readable.)

Here are two other generalizations.

Corollary (to CMT). The uniqueness and convergence conclusions of the CMT also hold if T satisfies the weaker condition

$$\|(1-\beta)(\boldsymbol{x}-\boldsymbol{y}) + \beta[T(\boldsymbol{x}) - T(\boldsymbol{y})]\| \le \alpha \|\boldsymbol{x}-\boldsymbol{y}\|, \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in S,$$

for any $\beta \neq 0$ and some $\alpha \in [0, 1)$, provided we use the following **damped** or **under-relaxed** iteration:

$$\boldsymbol{x}_{n+1} = (1-\beta)\boldsymbol{x}_n + \beta T(\boldsymbol{x}_n).$$

Proof. Let $T_\beta(\boldsymbol{x}) \triangleq (1-\beta)\boldsymbol{x} + \beta T(\boldsymbol{x})$, which is a contraction mapping with the same fixed point(s) as T for $\beta \neq 0$. □

Example. Consider $T(\boldsymbol{x}) = \boldsymbol{x} + 3(\boldsymbol{z} - \boldsymbol{x})$. This has a unique fixed point, but it is not a contraction mapping, so the ordinary successive approximation algorithm is inapplicable. However

$$(1-\beta)(\boldsymbol{x}-\boldsymbol{y}) + \beta[T(\boldsymbol{x}) - T(\boldsymbol{y})] = (1-\beta)(\boldsymbol{x}-\boldsymbol{y}) + \beta[\boldsymbol{x} + 3(\boldsymbol{z}-\boldsymbol{x}) - \boldsymbol{y} - 3(\boldsymbol{z}-\boldsymbol{x})] = (1-3\beta)(\boldsymbol{x}-\boldsymbol{y}),$$

so for $\beta \in (0, 1/3)$, the above under-relaxed iteration will converge to the unique fixed point.

Exercise. (Text 10.2, p. 308)
If S is a compact subset of a normed space $(\mathcal{X}, \|\cdot\|)$ and T is a mapping from S into S such that

$$\|T(\boldsymbol{x}) - T(\boldsymbol{y})\| < \|\boldsymbol{x}-\boldsymbol{y}\|, \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in S \text{ and } \boldsymbol{x} \neq \boldsymbol{y},$$

then T has a unique fixed point and the method of successive approximations converges to that fixed point.

In this version, we have weakened the requirements on T, but strengthened the requirements on S.

(The book's problem states "Banach space," but in fact "normed space" is sufficient.)

Application: differential equations

When considering ordinary differential equations (ODEs) of the form

$$\dot{x}(t) = h(x(t), t),$$

on $[a, b]$, where $x(a) = x_0$ is prespecified.

A key question is whether the solution is **unique** on $[a, b]$.
The answer depends, of course, on h.

Suppose h satisfies the following **Lipschitz condition** (of order one) for $t \in [a, b]$:

$$|h(x, t) - h(y, t)| \le M|x - y|, \qquad \forall x, y \in \mathbb{R},$$

where M is a finite constant (independent of t, x and y). Under this assumption, we can apply the CMT to show:
- the ODE has a solution, and
- that solution is unique.

The ODE is equivalent to

$$x(t) = x(a) + \int_a^t h(x(s), s) \, \mathrm{d}s,$$

so we consider the space $C[a, b]$ (with the usual $\|\cdot\|_\infty$ norm) and define $T : \mathcal{X} \to \mathcal{X}$ by

$$T(\boldsymbol{x}) = x_0 + \int_a^t h(x(s), s) \, \mathrm{d}s \, .$$

Then

$$\begin{aligned}
\|T(\boldsymbol{x}) - T(\boldsymbol{y})\|_\infty &= \max_{t \in [a,b]} \left| \int_a^t h(x(s), s) - h(y(s), s) \, \mathrm{d}s \right| \le \max_{t \in [a,b]} \int_a^t |h(x(s), s) - h(y(s), s) \, \mathrm{d}s| \\
&\le \max_{t \in [a,b]} \int_a^t M \, |x(s) - y(s)| \, \mathrm{d}s \le \max_{t \in [a,b]} \int_a^t M \, \|\boldsymbol{x} - \boldsymbol{y}\|_\infty \, \mathrm{d}s = (b-a)M \, \|\boldsymbol{x} - \boldsymbol{y}\|_\infty \, .
\end{aligned}$$

So if $M \le 1/(b - a)$, then T is a contraction mapping (on $\mathcal{X}$, which is inherently closed), so the CMT applies.

This is only a little exciting, since that constraint on M may be limiting. However,

$$\|T(\boldsymbol{x}) - T(\boldsymbol{y})\|_\infty \le (b - a)M \, \|\boldsymbol{x} - \boldsymbol{y}\|_\infty$$

shows that T is **continuous**, so if we could only show that T^m is a contraction, then we could apply our 2nd version of the CMT.

By a simple induction argument, one can show

$$\|T^m(\boldsymbol{x}) - T^m(\boldsymbol{y})\|_\infty \le \frac{M^m (b-a)^m}{m!} \, \|\boldsymbol{x} - \boldsymbol{y}\|_\infty \, ,$$

which can be made arbitrarily small by taking m sufficiently large. So T^m is a contraction, and hence
- T has a unique fixed point (that by construction is a solution to the ODE).
- In principle we could apply the method of successive approximations to find that solution.

Application: signal deconvolution ***overhead***

The **deconvolution** problem is to find an unknown signal $x[k]$ given the observed signal $z[k]$ under the convolution model:

$$z[k] = (x * h)[k] = \sum_{l=0}^{k} x[k-l]\, h[l], \qquad n = 0, 1, \ldots,$$

for a known **impulse response function** $h[k]$. We assume that[2] $h \in \ell_1$.
We also assume that the given $\boldsymbol{z}$ is in ℓ_∞, and that the corresponding $\boldsymbol{x}$ is also in ℓ_∞.

- This is a signal processing analog of the integral equation discussed in [4, Example 2, p. 275].
- For simplicity we consider a **causal** impulse response and a causal input signal, but the concepts generalize.

To yield a form amenable to the CMT, we subtract both sides from $x[k]$ and then rearrange as follows:

$$x[k] = x[k] + z[k] - (x * h)[k] = z[k] + (g * x)[k] \text{ where } g[k] \triangleq \delta[k] - h[k] = \begin{cases} -h[k], & k \neq 0 \\ 1 - h[0], & k = 0, \end{cases} \tag{10-1}$$

and $\delta[k]$ is the **Kronecker impulse function** which is unity for $k = 0$ and zero elsewhere. Clearly $g \in \ell_1$ since $h \in \ell_1$.

The form of (10-1) suggests that we try to use the following mapping: $T(\boldsymbol{x}) = z + g * x$.
We first need to pick an appropriate function space $\mathcal{X}$, show that $T : S \to S$ for some complete subset of $\mathcal{X}$.
Then we must show that T is a contraction mapping on S.
Let us try the Banach space ℓ_∞. If $\boldsymbol{y} = T_0(\boldsymbol{x})$ where $y[k] = \sum_{l=0}^{k} x[k-l]\, g[l]$ then

$$|y[k]| = \left| \sum_{l=0}^{k} x[k-l]\, g[l] \right| \leq \|x\|_\infty \|g\|_1 .$$

Thus $\|T_0(x)\|_\infty \leq \|x\|_\infty \|g\|_1$, and hence $\|T(x)\|_\infty \leq \|z\|_\infty + \|x\|_\infty \|g\|_1$, so we conclude that $T : \ell_\infty \to \ell_\infty$.

When is this mapping T a contraction on ℓ_∞?

$$\begin{aligned} \|T(\boldsymbol{x}) - T(\boldsymbol{y})\|_\infty &= \|(z + g * x) - (z + g * y)\|_\infty = \|g * (x - y)\|_\infty = \sup_k \left| \sum_{l=0}^{k} (x[k-l] - y[k-l])\, g[l] \right| \\ &\leq \sup_k \sum_{l=0}^{k} |g[l]|\, |x[k-l] - y[k-l]| \leq \sum_{l=0}^{k} |g[l]|\, \|x - y\|_\infty \leq \|g\|_1 \|x - y\|_\infty . \end{aligned}$$

So T is a contraction if $\|g\|_1 < 1$, in which case we can apply the method of successive approximations and "find" (asympototically) the unique fixed point of T, thereby solving the original deconvolution problem.

Unfortunately, $\|g\|_1 < 1$ holds only for a *very* restrictive class of impulse response functions $h[k]$, namely, those that are quite close to the impulse function $\delta[k]$, *i.e.*, $\|\delta - h\|_1 < 1$.

It is unsurprising that the class is restrictive, since deconvolution is possible only if there are no zeros in the frequency response, and is numerically stable only if the frequency response is bounded well away from zero.

. . . .

What about the under-relaxed version?

$$\begin{aligned} \|(1-\beta)(\boldsymbol{x} - \boldsymbol{y}) + \beta\, [T(\boldsymbol{x}) - T(\boldsymbol{y})]\|_\infty &= \|(1-\beta)(\boldsymbol{x} - \boldsymbol{y}) + \beta\, [g * (\boldsymbol{x} - \boldsymbol{y})]\|_\infty = \|(\boldsymbol{\delta} - \beta \boldsymbol{h}) * (\boldsymbol{x} - \boldsymbol{y})\|_\infty \\ &\leq \|\boldsymbol{\delta} - \beta \boldsymbol{h}\|_1 \|\boldsymbol{x} - \boldsymbol{y}\|_\infty . \end{aligned}$$

So it suffices to have $\|\boldsymbol{\delta} - \beta \boldsymbol{h}\|_1 < 1$. For that to hold, we must have $|h[0]| > \sum_{k \neq 0} |h[k]|$. This is still a restrictive class of impulse response functions. Within this class, the best choice of β is $\beta = \sum_{k \neq 0} |h[k]| \,/\, |h[0]|$.

. . . .

What about other mappings? Such as $T(x) = x + f * (z - h * x)$ where f is some filter.
What are sufficient conditions on f, h for this to satisfy the conditions of the under-relaxed version?

What if we are given that $z \in \ell_1$? **??**

[2] In discrete-time systems analysis this is equivalent to assuming that the system with impulse response $h[k]$ is bounded-input, bounded-output (BIBO) **stable**.

---

1. P. Enflo. A counterexample to the approximation problem in Banach spaces. *Acta Math*, 130:309–17, 1973.
2. I. J. Maddox. *Elements of functional analysis*. Cambridge, 2 edition, 1988.
3. A. W. Naylor and G. R. Sell. *Linear operator theory in engineering and science*. Springer-Verlag, New York, 2 edition, 1982.
4. D. G. Luenberger. *Optimization by vector space methods*. Wiley, New York, 1969.
5. J. Schauder. Zur theorie stetiger abbildungen in funktionenrumen. *Math. Zeitsch.*, 26:47–65, 1927.
6. L. Grafakos. *Classical and modern Fourier analysis*. Pearson, NJ, 2004.
7. P. P. Vaidyanathan. Generalizations of the sampling theorem: Seven decades after Nyquist. *IEEE Tr. Circ. Sys. I, Fundamental theory and applications*, 48(9):1094–109, September 2001.
8. A. M. Ostrowski. *Solution of equations in Euclidian and Banach spaces*. Academic, 3 edition, 1973.
9. R. R. Meyer. Sufficient conditions for the convergence of monotonic mathematical programming algorithms. *J. Comput. System. Sci.*, 12(1):108–21, 1976.
10. M. Rosenlicht. *Introduction to analysis*. Dover, New York, 1985.
11. A. R. De Pierro. On the relation between the ISRA and the EM algorithm for positron emission tomography. *IEEE Tr. Med. Im.*, 12(2):328–33, June 1993.
12. A. R. De Pierro. On the convergence of the iterative image space reconstruction algorithm for volume ECT. *IEEE Tr. Med. Im.*, 6(2):174–175, June 1987.
13. A. R. De Pierro. Unified approach to regularized maximum likelihood estimation in computed tomography. In *Proc. SPIE 3171, Comp. Exper. and Num. Meth. for Solving Ill-Posed Inv. Imaging Problems: Med. and Nonmed. Appl.*, pages 218–23, 1997.
14. J. A. Fessler. Grouped coordinate descent algorithms for robust edge-preserving image restoration. In *Proc. SPIE 3170, Im. Recon. and Restor. II*, pages 184–94, 1997.
15. A. R. De Pierro. A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. *IEEE Tr. Med. Im.*, 14(1):132–137, March 1995.
16. J. A. Fessler and A. O. Hero. Penalized maximum-likelihood image reconstruction using space-alternating generalized EM algorithms. *IEEE Tr. Im. Proc.*, 4(10):1417–29, October 1995.
17. M. W. Jacobson and J. A. Fessler. Properties of optimization transfer algorithms on convex feasible sets. *SIAM J. Optim.*, 2003. Submitted.
18. P. L. Combettes and H. J. Trussell. Method of successive projections for finding a common point of sets in metric spaces. *J. Optim. Theory Appl.*, 67(3):487–507, December 1990.
19. F. Deutsch. The convexity of Chebyshev sets in Hilbert space. In A. Yanushauskas Th. M. Rassias, H. M. Srivastava, editor, *Topics in polynomials of one and several variables and their applications*, pages 143–50. World Sci. Publishing, River Edge, NJ, 1993.
20. M. Jiang. On Johnson's example of a nonconvex Chebyshev set. *J. Approx. Theory*, 74(2):152–8, August 1993.
21. V. S. Balaganskii and L. P. Vlasov. The problem of convexity of Chebyshev sets. *Russian Mathematical Surveys*, 51(6):1127–90, November 1996.
22. V. Kanellopoulos. On the convexity of the weakly compact Chebyshev sets in Banach spaces. *Israel Journal of Mathematics*, 117:61–9, 2000.
23. A. R. Alimov. On the structure of the complements of Chebyshev sets. *Functional Analysis and Its Applications*, 35(3):176–82, July 2001.
24. M. Vetterli and J. Kovacevic. *Wavelets and subband coding*. Prentice-Hall, New York, 1995.
25. D. C. Youla and H. Webb. Image restoration by the method of convex projections: Part I—Theory. *IEEE Tr. Med. Im.*, 1(2):81–94, October 1982.