

STOPPING RULES FOR A NONNEGATIVELY CONSTRAINED ITERATIVE METHOD FOR ILL-POSED POISSON IMAGING PROBLEMS *

JOHNATHAN M. BARDSLEY †

*Department of Mathematical Sciences, University of Montana, Missoula, MT, 59812, USA.
email: bardsleyj@mso.umt.edu*

Abstract.

Image data is often collected by a charge coupled device (CCD) camera. CCD camera noise is known to be well-modeled by a Poisson distribution. If this is taken into account, the negative-log of the Poisson likelihood is the resulting data-fidelity function. We derive, via a Taylor series argument, a weighted least squares approximation of the negative-log of the Poisson likelihood function. The image deblurring algorithm of interest is then applied to the problem of minimizing this weighted least squares function subject to a nonnegativity constraint. Our objective in this paper is the development of stopping rules for this algorithm. We present three stopping rules and then test them on data generated using two different true images and an accurate CCD camera noise model. The results indicate that each of the three stopping rules is effective.

AMS subject classification (2000): 65F20, 65F30.

Key words: iterative methods, image reconstruction, regularization, statistical methods.

1 Introduction.

The following problem is very common in imaging science: given a blurred, noisy $N \times N$ image array \mathbf{z} , obtain an estimate of the underlying $N \times N$ true object array $\mathbf{u}_{\text{exact}}$ by approximately solving a linear system of the form

$$(1.1) \quad \mathbf{z} = \mathbf{A}\mathbf{u},$$

where \mathbf{z} has been column-stacked so that it is $N^2 \times 1$, and \mathbf{A} is a known $N^2 \times N^2$ ill-conditioned matrix. For the remainder of the manuscript, we will use the convention $n \stackrel{\text{def}}{=} N^2$.

The statistical model that is often used to simulate data noise in image processing is

$$(1.2) \quad \mathbf{z} = \mathbf{A}\mathbf{u}_e + \mathbf{e},$$

*Received

†This work was partially supported by the NSF under grant DMS-0504325

where \mathbf{u}_e is the true discrete image, and $\mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$, where \mathbf{I}_n is the $n \times n$ identity matrix and σ^2 the noise variance; that is, \mathbf{e} is an independent and identically distributed (i.i.d.) Gaussian random vector.

However, in both astronomical and medical imaging applications, the image \mathbf{z} is collected by a charge couple device (CCD) camera, which introduces noise that is not i.i.d. Gaussian. A statistical model for CCD camera noise that is significantly more accurate than (1.2) is

$$(1.3) \quad \mathbf{z} + \boldsymbol{\sigma}^2 = \text{Poiss}(\mathbf{A}\mathbf{u}_e + \boldsymbol{\gamma} + \boldsymbol{\sigma}^2).$$

Here $\boldsymbol{\gamma} = \gamma \mathbf{1}$ and $\boldsymbol{\sigma}^2 = \sigma^2 \mathbf{1}$, where $\mathbf{1}$ is the $n \times 1$ constant 1-vector, γ corresponds to the background intensity of the image being viewed and σ^2 is the variance of the instrument readout noise; and $\text{Poiss}(\boldsymbol{\lambda})$ is a Poisson random vector with Poisson parameter vector $\boldsymbol{\lambda}$ (see [1, 15] for more details regarding this statistical model).

The probability density function for \mathbf{z} then has the form

$$(1.4) \quad p_{\mathbf{z}}(\mathbf{z}; \mathbf{u}) := \prod_{i=1}^n \frac{([\mathbf{A}\mathbf{u}]_i + \gamma + \sigma^2)^{z_i + \sigma^2} \exp[-([\mathbf{A}\mathbf{u}]_i + \gamma + \sigma^2)]}{z_i + \sigma^2}.$$

We note that since Poisson random variables take on only discrete values, $p_{\mathbf{z}}(\mathbf{z}; \mathbf{u})$ should, in theory, be defined only for $\mathbf{z} \in \mathbb{Z}_+^n$. However to ease in both analysis and computation, we will treat $p_{\mathbf{z}}$ as a probability density function defined on $\mathbb{R}_+^n \cup \{0\}$.

Given image data \mathbf{z} arising from model (1.3), the maximum likelihood estimate of \mathbf{u}_e is obtained by maximizing $p_{\mathbf{z}}(\mathbf{z}; \mathbf{u})$ with respect to \mathbf{u} , or, more commonly, by minimizing $-\ln p_{\mathbf{z}}(\mathbf{z}; \mathbf{u})$ subject to the constraint $\mathbf{u} \geq \mathbf{0}$; that is, by solving

$$(1.5) \quad \mathbf{u}_{\text{ML}} = \arg \min_{\mathbf{u} \geq \mathbf{0}} T(\mathbf{u}; \mathbf{z}),$$

where $\mathbf{u} \geq \mathbf{0}$ means $u_i \geq 0$ for $i = 1, 2, \dots, n$, and

$$(1.6) \quad T(\mathbf{u}; \mathbf{z}) \stackrel{\text{def}}{=} \sum_{i=1}^n \{([\mathbf{A}\mathbf{u}]_i + \gamma + \sigma^2) - (z_i + \sigma^2) \ln([\mathbf{A}\mathbf{u}]_i + \gamma + \sigma^2)\}.$$

For given data \mathbf{z} , blurring matrix \mathbf{A} , and parameters γ and σ^2 , we will call (1.5), (1.6) a *Poisson imaging problem*.

As is the case for least squares estimation, due to the ill-conditioning in \mathbf{A} , solutions of (1.5) are unstable with respect to the noise in \mathbf{z} . Thus regularization is required. One way that this can be accomplished is by truncating an iterative method applied to (1.5) [3]. When this approach is taken, the choice of stopping iteration becomes extremely important and is akin to the choice of regularization parameter in the standard Tikhonov approach to regularization [3, 17].

Previous work has been done in this direction in relation to the widely used Richardson-Lucy (RL) algorithm—an iterative method for solving (1.5) [6, 13]—by researchers in positron emission tomography (PET) [5, 10, 12]. Some work

in astronomical imaging on stopping rules for RL was also done in [11], however no concrete stopping rules were given.

It was shown in [1] that the long-named *covariance-preconditioned, modified residual norm steepest descent method*—which we give the unfortunately long acronym CPMRNSD—has better convergence properties than RL when used on image data corrupted by Poisson noise. However stopping rules for CPMRNSD iterations have not been introduced. The development of such rules is the focus of this paper. With these rules in hand, and given the favorable comparisons between CPMRNSD and RL found in [1], we feel that CPMRNSD is a viable alternative to RL for use on Poisson imaging problems.

The paper is organized as follows. We begin in Section 2 with a derivation of a quadratic Taylor series approximation of $T(\mathbf{u}; \mathbf{z})$, from which the CPMRNSD method and its stopping rules are derived. We then present CPMRNSD in Section 3 and its stopping rules in Section 4. Numerical experiments are then the subject of Section 5. And finally, we end with conclusions in Section 6.

2 A Weighted Least Squares Approximation of $T(\mathbf{u}, \mathbf{z})$

Existing methods for estimating the regularization parameter in the least squares case [3, 17] can not be directly applied to our problem. However, in this section we will show that T (defined in (1.5)) is well approximated near the exact solution \mathbf{u}_e and exact data $\mathbf{z}_e \stackrel{\text{def}}{=} \mathbf{A}\mathbf{u}_e + \boldsymbol{\gamma}$ by a certain weighted least squares function.

First, we compute various derivatives of T . The gradient and Hessian of T with respect to \mathbf{u} are given by

$$(2.1) \quad \nabla_{\mathbf{u}} T(\mathbf{u}; \mathbf{z}) = \mathbf{A}^T \left(\frac{\mathbf{A}\mathbf{u} - (\mathbf{z} - \boldsymbol{\gamma})}{\mathbf{A}\mathbf{u} + \boldsymbol{\gamma} + \boldsymbol{\sigma}^2} \right),$$

$$(2.2) \quad \nabla_{\mathbf{u}\mathbf{u}}^2 T(\mathbf{u}; \mathbf{z}) = \mathbf{A}^T \text{diag} \left(\frac{\mathbf{z} + \boldsymbol{\sigma}^2}{(\mathbf{A}\mathbf{u} + \boldsymbol{\gamma} + \boldsymbol{\sigma}^2)^2} \right) \mathbf{A},$$

where division – here and for the remainder of the manuscript – is computed component-wise. The gradient and Hessian of T with respect to \mathbf{z} are given by

$$(2.3) \quad \nabla_{\mathbf{z}} T(\mathbf{u}; \mathbf{z}) = -\log(\mathbf{A}\mathbf{u} + \boldsymbol{\gamma} + \boldsymbol{\sigma}^2),$$

$$(2.4) \quad \nabla_{\mathbf{z}\mathbf{z}}^2 T(\mathbf{u}; \mathbf{z}) = \mathbf{0}.$$

The second order mixed partial derivatives of T are given by

$$(2.5) \quad \nabla_{\mathbf{u}\mathbf{z}}^2 T(\mathbf{u}; \mathbf{z}) = -\mathbf{A}^T \text{diag} \left(\frac{\mathbf{1}}{\mathbf{A}\mathbf{u} + \boldsymbol{\gamma} + \boldsymbol{\sigma}^2} \right),$$

$$(2.6) \quad \nabla_{\mathbf{z}\mathbf{u}}^2 T(\mathbf{u}; \mathbf{z}) = -\text{diag} \left(\frac{\mathbf{1}}{\mathbf{A}\mathbf{u} + \boldsymbol{\gamma} + \boldsymbol{\sigma}^2} \right) \mathbf{A}.$$

Now, let \mathbf{u}_e be the exact object and $\mathbf{z}_e \stackrel{\text{def}}{=} \mathbf{A}\mathbf{u}_e + \boldsymbol{\gamma}$ the background shifted exact data. Then, letting $\mathbf{k} = \mathbf{z} - \mathbf{z}_e$ and $\mathbf{h} = \mathbf{u} - \mathbf{u}_e$ and expanding T in a

Taylor series about \mathbf{u}_e and \mathbf{z}_e , we obtain from (2.1)-(2.6)

$$\begin{aligned} T(\mathbf{u}; \mathbf{z}) &= T(\mathbf{u}_e + \mathbf{h}; \mathbf{z}_e + \mathbf{k}), \\ &= T(\mathbf{u}_e; \mathbf{z}_e) - \mathbf{k}^T \nabla_{\mathbf{z}} T(\mathbf{u}_e; \mathbf{z}_e) + \frac{1}{2} \mathbf{h}^T \nabla_{\mathbf{u}\mathbf{u}}^2 T(\mathbf{u}_e; \mathbf{z}_e) \mathbf{h} + \frac{1}{2} \mathbf{k}^T \nabla_{\mathbf{z}\mathbf{z}}^2 T(\mathbf{u}_e; \mathbf{z}_e) \mathbf{k}, \\ &\quad + \frac{1}{2} \mathbf{h}^T \nabla_{\mathbf{z}\mathbf{u}}^2 T(\mathbf{u}_e; \mathbf{z}_e) \mathbf{k} + \mathcal{O}(\|\mathbf{h}\|_2^3, \|\mathbf{k}\|_2^3) \\ &= T(\mathbf{u}_e; \mathbf{z}) + \frac{1}{2} (\mathbf{A}\mathbf{u} - (\mathbf{z} - \gamma))^T \text{diag} \left(\frac{\mathbf{1}}{\mathbf{A}\mathbf{u}_e + \gamma + \sigma^2} \right) (\mathbf{A}\mathbf{u} - (\mathbf{z} - \gamma)) \\ &\quad + \mathcal{O}(\|\mathbf{h}\|_2^3, \|\mathbf{k}\|_2^3). \end{aligned}$$

Thus the quadratic Taylor series approximation of $T(\mathbf{u}; \mathbf{z})$ about the points $(\mathbf{u}_e; \mathbf{z}_e)$ is given by

$$(2.7) \quad T(\mathbf{u}_e; \mathbf{z}) + \frac{1}{2} (\mathbf{A}\mathbf{u} - (\mathbf{z} - \gamma))^T \text{diag} \left(\frac{\mathbf{1}}{\mathbf{A}\mathbf{u}_e + \gamma + \sigma^2} \right) (\mathbf{A}\mathbf{u} - (\mathbf{z} - \gamma)).$$

This Taylor series connection between the negative-log of the Poisson likelihood and (2.7) has, to the author's knowledge, not been noted elsewhere.

It is important to emphasize that the quadratic approximation (2.7) of T will be accurate provided $\|\mathbf{h}\|$ and $\|\mathbf{k}\|$ are small *relative to* $\|\mathbf{u}_e\|$ and $\|\mathbf{z}_e\|$. Since this will hold in practice for typical data \mathbf{z} and reasonable approximations of \mathbf{u} , we can expect that (2.7) will be an accurate approximation of $T(\mathbf{u}; \mathbf{z})$ in a region of $(\mathbf{u}_e; \mathbf{z}_e)$ that is not restrictively small.

In practice, if this approximation is to be used, $\mathbf{A}\mathbf{u}_e$ must in turn be approximated. The natural choice is to use $\mathbf{z} + \sigma^2$ instead since $E(\mathbf{z}) = \mathbf{A}\mathbf{u}_e + \gamma$, where E is the expected value function. This yields the following weighted least squares approximation of the Poisson likelihood function

$$(2.8) \quad T_{\text{wls}}(\mathbf{u}; \mathbf{z}) \stackrel{\text{def}}{=} \frac{1}{2} \|\mathbf{C}^{-1/2} (\mathbf{A}\mathbf{u} - (\mathbf{z} - \gamma))\|^2, \quad \text{where } \mathbf{C} = \text{diag}(\mathbf{z} + \sigma^2).$$

Note that the constant term $T(\mathbf{u}_e; \mathbf{z})$ has been dropped since it does not effect the computations.

In a large number of applications in which Poisson data is analyzed, T is approximated by a weighted least squares function in (1.5) (see, e.g., [2, 14]). Our analysis above suggests that T_{wls} is the natural choice, which yields the following approximation of (1.5):

$$(2.9) \quad \mathbf{u}_{\text{WLS}} = \arg \min_{\mathbf{u} \geq \mathbf{0}} T_{\text{wls}}(\mathbf{u}; \mathbf{z}).$$

We also view (2.9) as a Poisson imaging problem.

3 The CPMRNSD Algorithm

The CPMRNSD algorithm of [1] is obtained by applying the MRNSD algorithm of [4, 9] to (2.9). It has the form

$$(3.1) \quad \mathbf{u}_{k+1} = \mathbf{u}_k - \tau_k \mathbf{u}_k \odot \mathbf{A}^T \mathbf{C}^{-1} (\mathbf{A}\mathbf{u}_k - (\mathbf{z} - \gamma)),$$

where “ \odot ” denotes Hadamard (component-wise) multiplication, and the line search parameter τ_k in (3.1) is given by

$$\tau_k = \min\{\tau_{\text{uc}}, \tau_{\text{bd}}\}.$$

Here, if $\mathbf{v}_k = \mathbf{u}_k \odot \nabla J(\mathbf{u}_k)$,

$$(3.2) \quad \tau_{\text{uc}} = \frac{\langle \mathbf{v}_k, \nabla J(\mathbf{u}_k) \rangle_2}{\langle \mathbf{v}_k, \mathbf{A}^T \mathbf{C}^{-1} \mathbf{A} \mathbf{v}_k \rangle_2},$$

and

$$\tau_{\text{bd}} = \min \{ [\mathbf{u}_k]_i / [\mathbf{v}_k]_i \mid [\mathbf{v}_k]_i > 0, [\mathbf{u}_k]_i \neq 0 \}.$$

4 Regularization Parameter Choice Methods

In this section, we present introduce three different stopping rules for CPM-RNSD.

4.1 The Discrepancy Principle Stopping Rule

From (2.7) and (2.8) we have

$$E(T(\mathbf{u}; \mathbf{z})) \approx T(\mathbf{u}_e; \mathbf{z}_e) + E(T_{\text{wls}}(\mathbf{u}; \mathbf{z})),$$

where E is the expected value function. Thus it is reasonable to say that acceptable iterates \mathbf{u}_k will be those for which

$$T_{\text{wls}}(\mathbf{u}_k; \mathbf{z}) \approx E(T_{\text{wls}}(\mathbf{u}_e; \mathbf{z})).$$

In order to obtain an estimate of $E(T_{\text{wls}}(\mathbf{u}_e; \mathbf{z}))$ we first approximate (1.3) as is often done in practice (see, e.g., [1, 2, 14]):

$$(4.1) \quad \mathbf{z} - \boldsymbol{\gamma} = \mathbf{A} \mathbf{u}_e + \mathbf{e},$$

where $\mathbf{e} \sim N(\mathbf{0}, \mathbf{C})$ with \mathbf{C} defined in (2.8). This in turn implies

$$(4.2) \quad \mathbf{r}(\mathbf{u}_e) \sim N(\mathbf{0}, \mathbf{I}_n),$$

where

$$(4.3) \quad \mathbf{r}(\mathbf{u}) \stackrel{\text{def}}{=} \mathbf{C}^{-1/2}(\mathbf{A} \mathbf{u} - (\mathbf{z} - \boldsymbol{\gamma})).$$

A standard statistical result then tells us that given (4.2),

$$(4.4) \quad \|\mathbf{r}(\mathbf{u}_e)\|_2^2 \sim \chi^2(n),$$

where $\chi^2(n)$ denotes the chi-squared distribution with n degrees of freedom, which has mean and variance n and $2n$, respectively. Hence, in practice we should have (keep in mind the approximations that have been made) $E(T_{\text{wls}}(\mathbf{u}_e; \mathbf{z})) \approx$

n . This motivates the following acceptability criterion for a CPMRNSD iterate \mathbf{u}_k :

$$2T_{\text{wls}}(\mathbf{u}_k; \mathbf{z}) \approx n,$$

More specifically, we can say that the appropriate iterates are those for which for which $\|\mathbf{r}(\mathbf{u}_k)\|^2$ lies within two standard deviations of n ; that is, if

$$(4.5) \quad n - 2\sqrt{2n} \leq 2T_{\text{wls}}(\mathbf{u}_k; \mathbf{z}) \leq n + 2\sqrt{2n}.$$

We note that this interval accounts for roughly 95% of the area under the $\chi^2(n)$ curve.

However, for the purpose of developing a stopping rule for CPMRNSD iterations, only the upper bound is important since in general $2T_{\text{wls}}(\mathbf{u}_k)$ is typically much larger than $n + 2\sqrt{2n}$ in early iterations and decreases with k . More generally, we suggest the stopping rule

$$(4.6) \quad \frac{2}{n} T_{\text{wls}}(\mathbf{u}_k; \mathbf{z}) \leq 1 + \epsilon_n,$$

where ϵ_n is taken to be, for example, 0, $\sqrt{2n}/n$, or $2\sqrt{2n}/n$. We note that $\epsilon_n = 0$ corresponds to Morozov's discrepancy principle [3, 7, 17].

The convergence properties of the algorithm being used will, in large part, determine the choice of ϵ_n , which can be treated as a separate parameter that the application scientist tunes based on her experience with the problem at hand. This gives the method some additional flexibility, which may be advantageous in practice. In choosing ϵ_n , one should not deviate too far from (4.5). We note that a similar stopping rule was proposed in [16] for regular least squares problems.

4.2 The Method of Generalized Cross Validation

In [10, 12] it is shown that the generalized cross validation (GCV) method [3, 17, 18] can be used to develop stopping rules for the steepest descent and RL iterations. Following their approach, we can develop additional stopping rules for CPMRNSD.

The GCV function for (2.8) at iteration k is defined by

$$(4.7) \quad \text{GCV}(k) = n \|\mathbf{C}^{-1/2}(\mathbf{A}\mathbf{u}_k - (\mathbf{z} - \boldsymbol{\gamma}))\|^2 / \text{trace}(\mathbf{I}_n - \mathbf{C}^{-1/2}\mathbf{A}\mathbf{A}_k)^2,$$

where \mathbf{A}_k is the iterative regularization matrix satisfying

$$\mathbf{A}_k(\mathbf{z} - \boldsymbol{\gamma}) = \mathbf{u}_k,$$

with \mathbf{u}_k the k th CPMRNSD iterate. The idea is then to stop CPMRNSD at the iteration k that results in an increase in the GCV function.

In order to evaluate $\text{GCV}(k)$ in practice $\text{trace}(\mathbf{I}_n - \mathbf{C}^{-1/2}\mathbf{A}\mathbf{A}_k)$ must be approximated. This can be accomplished using the Trace Lemma [17]: given $\mathbf{B} \in \mathbb{R}^{n \times n}$,

$$(4.8) \quad \mathbf{v} \sim N(\mathbf{0}, \mathbf{I}_n) \quad \text{implies} \quad E(\mathbf{v}^T \mathbf{B} \mathbf{v}) = \text{trace}(\mathbf{B}).$$

Thus given a realization \mathbf{v} from $N(\mathbf{0}, \mathbf{I}_n)$,

$$\text{trace}(\mathbf{I}_n - \mathbf{C}^{-1/2} \mathbf{A} \mathbf{A}_k) \approx \mathbf{v}^T \mathbf{v} - \mathbf{v}^T \mathbf{C}^{-1/2} \mathbf{A} \mathbf{A}_k \mathbf{v}.$$

It is shown in [17] that this approximation is accurate in practice.

Thus, if we know \mathbf{A}_k , we can efficiently estimate $\text{GCV}(k)$ at each iteration. However we don't know \mathbf{A}_k , nor do we want to compute it. Instead, we follow [10] and define

$$\mathbf{w}_k = \mathbf{A}_k \mathbf{v},$$

which yields the following approximation of (4.7):

$$(4.9) \quad \widetilde{\text{GCV}}(k) = n \|\mathbf{C}^{-1/2}(\mathbf{A} \mathbf{u}_k - (\mathbf{z} - \boldsymbol{\gamma}))\|^2 / (\mathbf{v}^T \mathbf{v} - \mathbf{v}^T \mathbf{C}^{-1/2} \mathbf{A} \mathbf{w}_k).$$

It remains, only, to compute \mathbf{w}_k . To do this, we note that if $\mathbf{z}_k \stackrel{\text{def}}{=} \mathbf{A} \mathbf{u}_k = \mathbf{A} \mathbf{A}_k (\mathbf{z} - \boldsymbol{\gamma})$, then

$$[\mathbf{A} \mathbf{A}_k]_{i\ell} = \frac{\partial [\mathbf{z}_k]_i}{\partial z_\ell},$$

and hence,

$$\begin{aligned} [\mathbf{A} \mathbf{w}_k]_i &= [\mathbf{A} \mathbf{A}_k \mathbf{v}]_i \\ &= \sum_{\ell} \frac{\partial [\mathbf{z}_k]_i}{\partial z_\ell} v_\ell \\ &= \sum_{\ell} \sum_j a_{ij} \frac{\partial [\mathbf{u}_k]_j}{\partial z_\ell} v_\ell \\ &= \sum_j a_{ij} \sum_{\ell} \frac{\partial [\mathbf{u}_k]_j}{\partial z_\ell} v_\ell. \end{aligned}$$

Thus

$$(4.10) \quad [\mathbf{w}_k]_i = \frac{\partial [\mathbf{u}_k]_i}{\partial z_\ell} v_\ell.$$

From (4.10), \mathbf{w}_k iterations for a large class of iterative methods can be developed. In [10], this is done for steepest descent, whereas in [12] it is done for RL. We now do it for CPMRNSD.

The \mathbf{w}_k iteration for CPMRNSD is derived following the general approach outlined in [10]. In particular, we component-wise differentiate both sides of (3.1) with respect to z_ℓ for $\ell = 1, \dots, n$ and multiply each component by v_ℓ . Then, using basic calculus and (4.10), one obtains the iteration

$$(4.11) \quad \begin{aligned} \mathbf{w}_{k+1} &= \mathbf{w}_k - \tau_k [\mathbf{w}_k \odot \mathbf{A}^T \mathbf{C}^{-1} (\mathbf{A} \mathbf{u}_k - (\mathbf{z} - \boldsymbol{\gamma})) \\ &\quad + \mathbf{u}_k \odot \mathbf{A}^T \mathbf{C}^{-1} (\mathbf{A} \mathbf{w}_k - \mathbf{v})], \end{aligned}$$

where $\mathbf{w}_0 = (\mathbf{u}_0 \odot \mathbf{v}) / (\mathbf{z} - \boldsymbol{\gamma})$.

Thus we immediately have a stopping rule for CPMRNSD iterations; namely, iterate (3.1) and (4.11) simultaneously, and stop iterations once $\widetilde{\text{GCV}}(k) > \widetilde{\text{GCV}}(k-1)$, where $\widetilde{\text{GCV}}(k)$ is defined in (4.7).

We note that this extends in the obvious way to the case where $\mathbf{C} = \mathbf{I}_n$ in (3.1), (4.7), and (4.11), thus yielding a stopping law for the MRNSD algorithm of [9]. This approach can also be applied to a number of other methods, such as conjugate gradient iterations, though the author's experiments suggests that it is less effective for unconstrained iterative methods.

4.3 An Alternative to GCV: the Unbiased Predictive Risk Estimator

An alternative to GCV that requires knowledge of the noise variance—which we have in some approximate sense (see (4.2))—is the unbiased predictive risk estimator (UPRE) [17]. Here the goal is to choose the iteration k for which $E\left(\frac{2}{n}T_{wls}(\mathbf{u}_k; \mathbf{z}_e)\right)$ —known as the *predictive risk*—is smallest. The UPRE function for (2.8) is

$$(4.12) \quad \text{UPRE}(k) = \frac{2}{n}T_{wls}(\mathbf{u}_k; \mathbf{z}) + \frac{2}{n} \text{trace} \left(\mathbf{C}^{-1/2} \mathbf{A} \mathbf{A}_k \right) - 1.$$

It is derived as in [17] for regular least squares problems and is an unbiased estimator of the predictive risk, hence its name. Following the same approach as for GCV, we can approximate (4.12) by

$$(4.13) \quad \widetilde{\text{UPRE}}(k) = \frac{2}{n}T_{wls}(\mathbf{u}_k; \mathbf{z}) + \frac{2}{n} \mathbf{v}^T \mathbf{C}^{-1/2} \mathbf{A} \mathbf{w}_k - 1,$$

where \mathbf{v} and \mathbf{w}_k are as above.

Thus we have a third stopping rule for CPMRNSD iterations; namely, iterate (3.1) and (4.11) simultaneously, and stop iterations once $\widetilde{\text{UPRE}}(k) > \widetilde{\text{UPRE}}(k-1)$, where $\widetilde{\text{UPRE}}(k)$ is defined in (4.13).

5 Numerical Results

In this section we present numerical results using two image restoration test problems. The first set of data was developed at the US Air Force Phillips Laboratory, Lasers and Imaging Directorate, Kirtland Air Force Base, New Mexico. The image is a computer simulation of a field experiment showing a satellite as taken from a ground based telescope. The true image has 256×256 pixels and is shown on the left in Fig. 5.1. Both the image and the blurring matrix \mathbf{A} used in our experiments are found in the *RestoreTools* image restoration package [8]. In fact, all of our numerical tests were performed within *RestoreTools*.

However, we imported the simulated star field data seen on the right in Figure 5.1, as well as the point spread function (PSF) that defines the blurring matrix \mathbf{A} . The PSF—supplied by Dr. Brent Ellerbroek of the Thirty Meter Telescope Project—is a simulation of the type of PSF that results from an adaptive optics telescope system. Such systems have PSFs that are much nearer to the diffraction limit than PSFs from systems without adaptive optics.

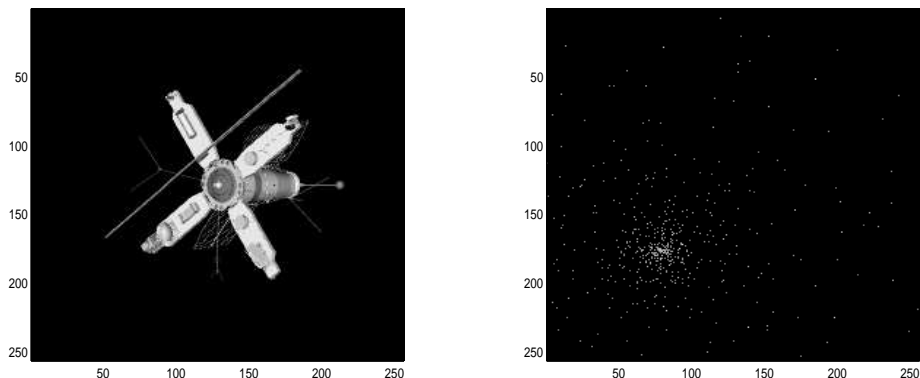


Figure 5.1: *Satellite test image on the left and star field test image (plotted on a log scale) on the right.*

To generate our data, we use the statistical model of [15], which has the form

$$(5.1) \quad \mathbf{z} = \text{Pois}(\mathbf{A}\mathbf{u}_e) + \text{Pois}(\boldsymbol{\gamma}) + N(\mathbf{0}, \sigma^2 \mathbf{I}_n),$$

We note that (1.3) is an approximation of (5.1) (see [1] for details). We generate Poisson noise using MATLAB's `poissrnd` function and choose the values $\gamma = 10$ and $\sigma = 5$, which are realistic values for these parameters. The blurred, noisy data with SNR=10 is plotted in Figure 5.2. We plot the star-field on a log-scale since its features are difficult to visualize otherwise.

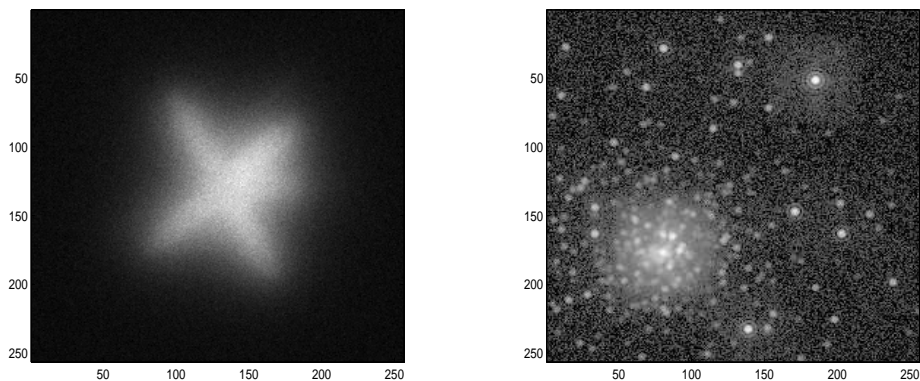


Figure 5.2: *Blurred, noisy satellite data on the left and star field data (plotted on a log scale) on the right, each with SNR=10.*

In our implementation of the CPMRNSD iteration, we chose an initial guess

of $\mathbf{u}_0 = \mathbf{1}$, the constant 1-vector. We also restarted the \mathbf{w}_k iteration (4.11) every 100 CPMRNSD iterations. We found that this enhanced the effectiveness of the stopping rules. This is accomplished by taking $\mathbf{w}_k = (\mathbf{u}_k \odot \mathbf{v}) / (\mathbf{z} - \gamma)$, where k is a multiple of 100. In order to ensure that an instability due to the restart itself isn't responsible for the increase in the GCV and UPRE curves, we allow for 10 CPMRNSD iterations before stopping iterations if either the GCV or UPRE functions increase.

In order to test our methods on a number of data sets, we vary the intensity of the true image \mathbf{u}_e in both test cases to obtain data containing noise at a number of different signal-to-noise ratios (SNRs), which for (5.1) is defined

$$SNR = \sqrt{\frac{\|\mathbf{A}\mathbf{u}_e\|^2}{\sum_{i=1}^n \{[\mathbf{A}\mathbf{u}_e]_i + \gamma + \gamma^2 + \sigma^2\}}}.$$

In particular, we perform experiments on each data set with SNR=1, 5, 10, and 30. We plot the relative error

$$\frac{\|\mathbf{u}_k - \mathbf{u}_e\|}{\|\mathbf{u}_e\|}$$

in each of our comparisons and mark the iteration at which the various stopping rules would stop iterations. We denote the discrepancy principle by DP and choose three different values of ϵ_n in (4.6), corresponding to 2, 1, and 0 standard deviations from the mean of $\chi^2(n)$: $1 + 2\sqrt{2n}/n$ (DP1), $1 + \sqrt{2n}/n$ (DP2), and 1 (DP3).

The results for the satellite data are plotted in Figure 5.3. The stopping iteration for the DP rules are not indicated in the first three plots because those thresholds were not met by the CPMRNSD iterations. The GCV and UPRE stopping rules are quite effective in each case, however for SNR=30 iterations are stopped somewhat early, though the result is still a high quality reconstruction.

The results for the star-field data are plotted in Figure 5.4. We note that though in the satellite test case we were able to obtain an increase in the relative error in a reasonable number of iterations, this was not possible with the star-field data for every SNR. In particular, only in the SNR=1 case were we able to obtain an increase in the relative error. For each SNR, the stopping rules advocated stopping iterations before the relative error had reached a minimum. However, in the star-field case, after a certain point, more iterations yields reconstructions that are visually indistinguishable when viewed using MATLAB's `imagesc` function, even though the relative error may be decreasing.

The results indicate that the stopping rules presented here for CPMRNSD iterations are useful. Moreover, given the fact that the GCV and UPRE stopping rules require no a priori choice of parameters, such as the ϵ_n parameter for DP, and that they seem to work well on a number of examples, we feel that they are superior to DP for the examples presented here. However, we note that DP will become more useful if the optimal value of ϵ_n is approximately known before iterations, which doesn't seem to be an impossibility in an applied setting. Finally, we note that GCV is a more flexible method than UPRE. In particular,

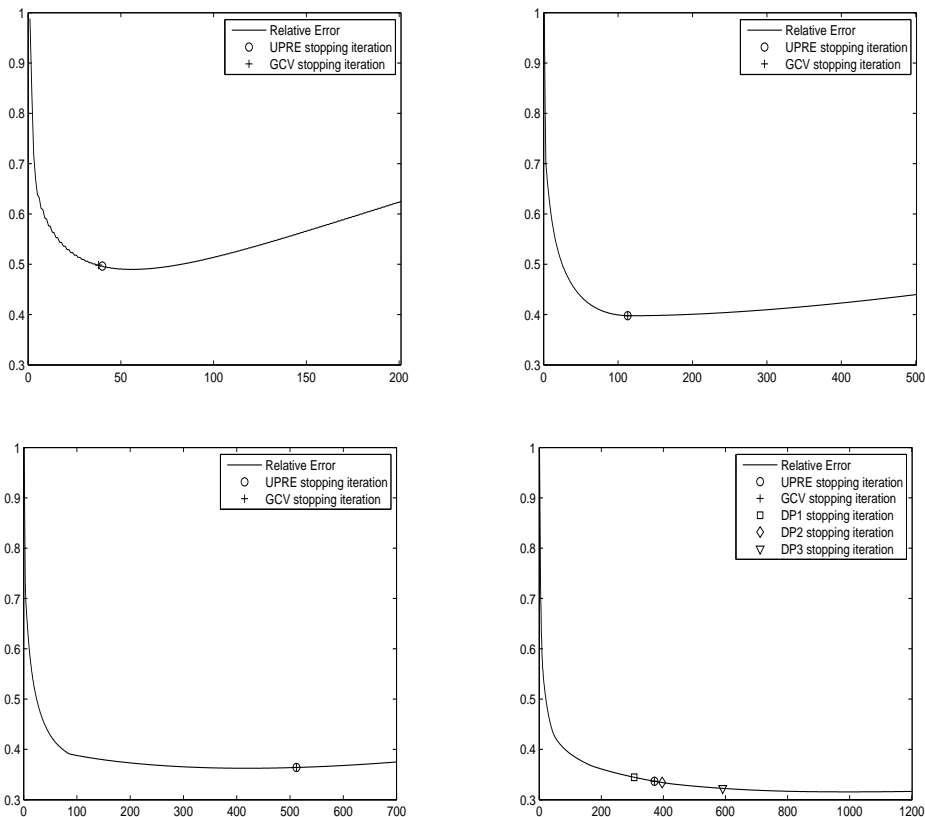


Figure 5.3: *Satellite Test Case.* Plot of relative error with stopping rules GCV, UPRE, DP1 ($\epsilon_n = 1 + 2\sqrt{2n/n}$), DP2 ($\epsilon_n = 1 + \sqrt{2n/n}$), DP3 ($\epsilon_n = 1$). The upper-left curve corresponds to SNR=1, the upper-right curve to SNR=5, the lower-left curve to SNR=10, and the lower-right curve to SNR=30.

it can be used on regular least squares problems in which the statistical model for the noise is not known. UPRE, on the other hand, requires knowledge of the noise variance.

It is interesting to observe that in all cases UPRE and GCV stopping rules stop iterations at the same—or very nearly the same—iteration. This suggests that the two methods are quite similar. This observation is supported by analysis found in [17], though there the UPRE method is shown to give better performance than GCV. Thus we think that it is important to present both methods here.

It is important to note that the stopping rule for GCV and UPRE will change somewhat with different realizations \mathbf{v} from $N(\mathbf{0}, \mathbf{I}_n)$ in (4.9) and (4.13). However, experiments indicate that though the stopping iteration does change with

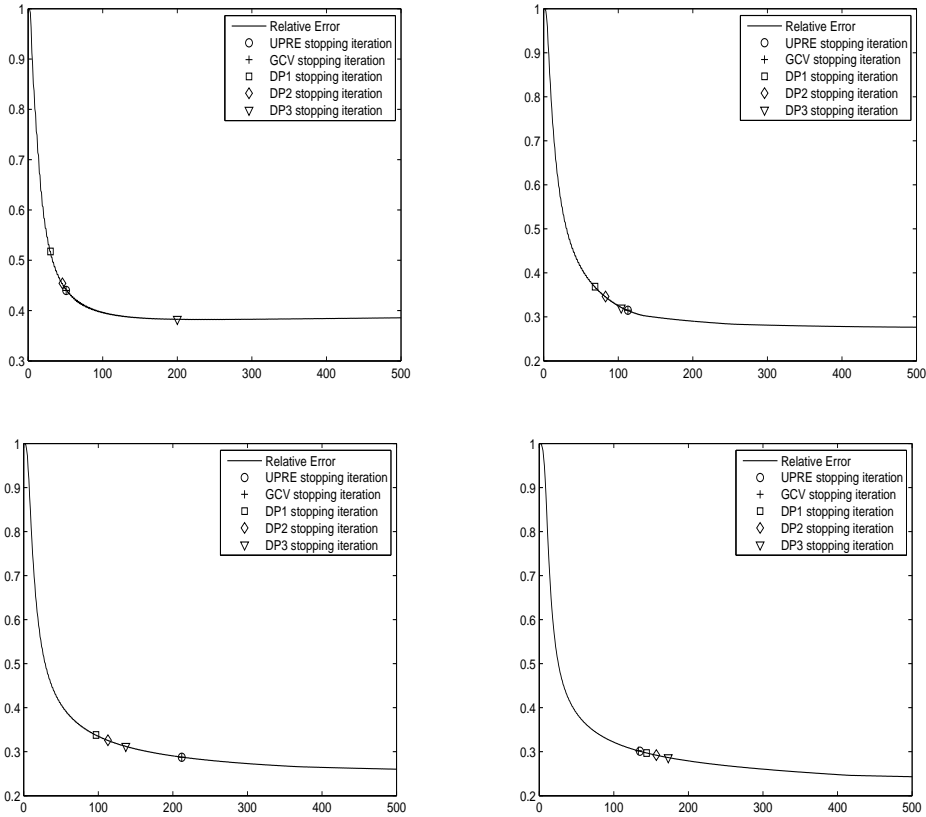


Figure 5.4: *Star Field Test Case*. Plot of relative error with stopping rules GCV, UPRE, DP1 ($\epsilon_n = 1 + 2\sqrt{2n}/n$), DP2 ($\epsilon_n = 1 + \sqrt{2n}/n$), DP3 ($\epsilon_n = 1$). The upper-left curve corresponds to SNR=1, the upper-right curve to SNR=5, the lower-left curve to SNR=10, and the lower-right curve to SNR=30.

\mathbf{v} , the GCV and UPRE rules are reasonably robust to these changes. A more stable stopping rule can be obtained by taking several realizations from $N(\mathbf{0}, \mathbf{I}_n)$, iterating (4.11) for each realization, and then averaging the corresponding \mathbf{w}_k 's. The increase in computational cost of this approach is clear, though the opportunity for parallelization is also evident.

6 Conclusions

Images collected with a CCD camera are corrupted with Poisson noise. When one takes this fact into account, the negative-log of the Poisson likelihood function T (defined in (1.4)) results. We have shown, via a Taylor series argument, that the weighted least squares approximation T_{wls} (defined in (2.8)) of T can

be expected to be accurate. Using this approximation in (1.5) yields the non-negatively constrained weighted least squares problem (2.9).

Applying the MRNSD algorithm [4, 9] to (2.9) yields the CPMRNSD algorithm of [1]. Our focus in this paper has been the development of stopping rules for CPMRNSD. We have presented three: the discrepancy principle (DP), generalized cross validation (GCV), and the unbiased predictive risk estimate (UPRE). Statistical motivation was given for DP, while in the case of both GCV and UPRE, an iterative implementation unique to CPMRNSD was provided.

In our numerical experiments, we performed tests using two separate synthetically generated objects: a satellite and a star-field. The data was generated using statistical model (5.1), which is the model for CCD camera noise statistics presented in [15]. By varying the intensity of the object, we were able to create data with SNRs of 1, 5, 10, and 30. Thus we tested the performance of the CPMRNSD iteration and its stopping rules on 8 different data sets. The results clearly indicate that the stopping rules are effective, making the CPMRNSD algorithm even more attractive as an iterative, Poisson imaging algorithm.

We concluded that the GCV and UPRE stopping rules were more effective than DP, however DP allows for the incorporation of a priori knowledge of the appropriate value of the parameter ϵ_n . The GCV and UPRE stopping rules were very similar for the examples considered here.

Finally, we note that the general approach presented here for implementing the GCV and UPRE stopping rules for CPMRNSD iterations should easily be extended to other methods, though initial experiments suggest that it seems to be more effective for nonnegatively constrained iterative methods.

Acknowledgements

The author would like to thank the students from his Fall Semester 2007 graduate course titled *Numerical Linear Algebra and Inverse Problems* for their helpful comments on the lectures based on the work in this paper and for their course final project work, which helped the author work out many of the early kinks in the methods.

REFERENCES

1. Johnathan M. Bardsley and James G. Nagy, *Covariance-Preconditioned Iterative Methods for Nonnegatively Constrained Image Reconstruction*, SIAM Journal on Matrix Analysis and Applications, 27 (4), 2006, pp. 1184-1198.
2. A. Grinvald and I. Steinberg, *On the Analysis of Fluorescence Decay Kinetics by the Method of Least-Squares*, Analytical Biochemistry, 59, 1974, pp.583-594.
3. P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, 1997.
4. Linda Kaufman, *Maximum Likelihood, Least Squares, and Penalized Least Squares for PET*, IEEE Transactions of Medical Imaging, 12(2), June 1993.
5. Jorge Llacer and Eugene Veklerov, *Feasible Images and Practical Stopping Rules for Iterative Algorithms in Emission Tomography*, IEEE Transactions on Medical Imaging, Vol. 8, No. 2, 1989, pp. 186-193.

6. L. B. Lucy, *An Iterative Technique for the Rectification of Observed Distributions*, The Astronomical Journal, **79** (1974), pp. 745-754.
7. V. A. Morozov, *On the solution of functional equations by the method of regularization*, Soviet Mathematics Doklady, vol. 7, 1966, pp. 414-417.
8. J. G. Nagy, K. Palmer, and L. Perrone, *Iterative Methods for Image Restoration: A Matlab Object Oriented Approach*, Numerical Algorithms, **36** (2003), pp. 73-93.
9. J. Nagy and Z. Strakoš, *Enforcing nonnegativity in image reconstruction algorithms*, Mathematical Modeling, Estimation, and Imaging, David C. Wilson, et.al., Eds., 4121 (2000), pg. 182-190.
10. Kevin M. Perry and Stanley J. Reeves, *Generalized Cross-Validation as a Stopping Rule for the Richardson-Lucy Algorithm*, in The Restoration of HST Images and Spectra II. R.J. Hanisch and R.L. White, Eds., Space Telescope Science Institute, 1994, pp. 97-103.
11. Sudhakar Prasad, *Statistical-information-based performance criteria for Richardson-Lucy image deblurring*, Journal of the Optical Society of America A, Vol. 19, No. 7, 2002, pp. 1286-1296.
12. Stanley J. Reeves, *Generalized Cross-Validation as a Stopping Rule for the Richardson-Lucy Algorithm*, International Journal of Imaging Systems and Technology, Vol. 6, 1995, pp. 387-391.
13. W. H. Richardson, *Bayesian-Based Iterative Method of Image Restoration*, Journal of the Optical Society of America **62** (1972), 55-59.
14. Bert Rust, *Truncating the Singular Value Decomposition for Ill-Posed Problems*, NISTIR 6131, U.S. Dept. of Commerce, July 1998
15. D. L. Snyder, A. M. Hammoud, and R. L. White, *Image recovery from data acquired with a charge-coupled-device camera*, Journal of the Optical Society of America A, **10** (1993), pp. 1014-1023.
16. H. J. Trussell, *Convergence criteria of iterative restoration methods*, IEEE Transactions on Acoustics, Speech, and Signal Processing, ASSP-31, 1983, pp. 129-136.
17. C. R. Vogel, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, 2002.
18. G. Wahba, *Practical approximate solutions to linear operator equations when the data are noisy*, SIAM Journal on Numerical Analysis, vol. 14, 1977, pp. 651-667.