Stochastic Process Estimation for AGN Reverberation (SPEAR)

Ying Zu, C.S. Kochanek and Bradley M. Peterson

November 5, 2010

1 Mathematical Description of the Method

1.1 Exponential Covariance

We start with a model process driving the continuum $s_c(t)$ that has a covariance between times t_i and t_j of

$$\langle s_c(t_i)s_c(t_j)\rangle = \sigma^2 \exp(-|t_i - t_j|/\tau). \tag{1}$$

Physically, the model corresponds to a random walk described by an amplitude $\sigma^2 = \hat{\sigma}^2\tau/2$ on long time scales with an exponential damping time scale τ, where $\hat{\sigma}$ and τ are used as our model parameters.

The light curve of the lines $s_l(t)$ is then

$$s_l(t) \equiv \int dt' g(t - t')s_c(t) \tag{2}$$

where $g(t-t')$ is the transfer function that determines the response of the lines to changes in the continuum. Since the lines and continuum are related by the transfer function, we can also determine the covariance between the line and continuum

$$\langle s_l(t_i)s_c(t_j)\rangle = \int dt' g(t_i - t')\langle s_c(t')s_c(t_j)\rangle, \tag{3}$$

between the line and itself

$$\langle s_l(t_i)s_l(t_j)\rangle = \int dt'dt'' g(t_i - t')g(t_j - t'')\langle s_c(t')s_c(t'')\rangle, \tag{4}$$

and between two different lines

$$\langle s_l(t_i)s_l'(t_j)\rangle = \int dt'dt'' g(t_i - t')g'(t_j - t'')\langle s_c(t')s_c(t'')\rangle. \tag{5}$$

1.2 Derivation of The Likelihood

Let $\mathbf{s}$ be a vector comprised of all the intrinsic signals of light curves, both line and continuum, and $S = \langle \mathbf{ss}\rangle$ be the covariance matrix between all the elements of $\mathbf{s}$. By definition, in Gaussian statistics, the probability of the intrinsic light curve is simply

$$P(\mathbf{s}) \propto |S|^{-1/2} \exp\left(-\frac{\mathbf{s}^T S^{-1}\mathbf{s}}{2}\right). \tag{6}$$

However, we do not measure the intrinsic light curve, but some realization of it, $\mathbf{y} = \mathbf{s} + \mathbf{n} + L\mathbf{q}$, in which there is measurement error $\mathbf{n}$, whose probability distribution is

$$P(\mathbf{n}) \propto |N|^{-1/2} \exp\left(-\frac{\mathbf{n}^T N^{-1} \mathbf{n}}{2}\right). \tag{7}$$

where $N = \langle \mathbf{n}\mathbf{n} \rangle$ is the covariance matrix of the noise. Note that nothing requires N to be diagonal, so there is no formal difficulty to including covariances in the noise between the lines and continuum. We have also allowed for the simultaneous fitting of a general trend defined by a response matrix L and a set of linear coefficients $\mathbf{q}$. In particular, we use this to fit and remove separate means from the light curves, and in this application, if solving a model with two light curves, L is a $2 \times N$ matrix with entries of $(1, 0)$ for the continuum data points and $(0, 1)$ for the line data points.

Given these definitions, the probability of the data $\mathbf{y}$ given the linear coefficients $\mathbf{q}$, the intrinsic light curves $\mathbf{s}$, and any other parameters of the model $\mathbf{p}$ is

$$\begin{aligned} P\left(\mathbf{y}\middle|\mathbf{q}, \mathbf{s}, \mathbf{p}\right) \;\propto\; & |SN|^{-1/2} \int d^n\mathbf{n}\, d^n\mathbf{s}\, \delta\left(\mathbf{y} - (\mathbf{s} + \mathbf{n} + L\mathbf{q})\right) \\ & \exp\left(-\frac{\mathbf{s}^T S^{-1}\mathbf{s} + \mathbf{n}^T N^{-1}\mathbf{n}}{2}\right). \end{aligned} \tag{8}$$

After evaluating the Dirac delta function, we "complete the squares" in the exponential with respect to both the unknown intrinsic source variability $\mathbf{s}$ and the linear coefficients $\mathbf{q}$. Define $\Re$ as the quadratic matrix term in the exponent,

$$\Re = \mathbf{s}^T S^{-1}\mathbf{s} + (\mathbf{y} - L\mathbf{q} - \mathbf{s})^T N^{-1}(\mathbf{y} - L\mathbf{q} - \mathbf{s}) \tag{9}$$

For simplicity, we set $\hat{\mathbf{y}} \equiv \mathbf{y} - L\mathbf{q}$, and complete the squares with respect to $\mathbf{q}$ later,

$$\begin{aligned} \Re &= \mathbf{s}^T S^{-1}\mathbf{s} + (\hat{\mathbf{y}} - \mathbf{s})^T N^{-1}(\hat{\mathbf{y}} - \mathbf{s}) \\ &= \mathbf{s}^T(S^{-1} + N^{-1})\mathbf{s} + \hat{\mathbf{y}}^T N^{-1}\hat{\mathbf{y}} - 2N^{-1}\hat{\mathbf{y}}\mathbf{s} \\ &= (\mathbf{s} - \hat{\mathbf{s}})^T(S^{-1} + N^{-1})(\mathbf{s} - \hat{\mathbf{s}}) - \hat{\mathbf{s}}^T(S^{-1} + N^{-1})\hat{\mathbf{s}} + \hat{\mathbf{y}}^T N^{-1}\hat{\mathbf{y}} \end{aligned} \tag{10}$$

where

$$\begin{aligned} \hat{\mathbf{s}} &= (S^{-1} + N^{-1})^{-1}N^{-1}\hat{\mathbf{y}} \\ &= S(S + N)^{-1}NN^{-1}\hat{\mathbf{y}} \\ &= S(S + N)^{-1}\hat{\mathbf{y}}. \end{aligned} \tag{11}$$

In the 2nd step of Eq. 11, we make use of the identity

$$(S^{-1} + N^{-1})^{-1} = S(S + N)^{-1}N = N(S + N)^{-1}S \tag{12}$$

Substituting Eq. (11) into the second term of Eq. (10),

$$\begin{aligned} \Re &= (\mathbf{s} - \hat{\mathbf{s}})^T(S^{-1} + N^{-1})(\mathbf{s} - \hat{\mathbf{s}}) - \hat{\mathbf{y}}^T(S + N)^{-1}SN^{-1}\hat{\mathbf{y}} + \hat{\mathbf{y}}^T N^{-1}\hat{\mathbf{y}} \\ &= (\mathbf{s} - \hat{\mathbf{s}})^T(S^{-1} + N^{-1})(\mathbf{s} - \hat{\mathbf{s}}) - \hat{\mathbf{y}}^T((S + N)^{-1}SN^{-1} - N^{-1})\hat{\mathbf{y}} \\ &= (\mathbf{s} - \hat{\mathbf{s}})^T(S^{-1} + N^{-1})(\mathbf{s} - \hat{\mathbf{s}}) + \hat{\mathbf{y}}^T(S + N)^{-1}\hat{\mathbf{y}}. \end{aligned} \tag{13}$$

Similarly, we can "complete the square" for the second term involving $\hat{\mathbf{y}}$ with respect to $\mathbf{q}$

$$\Re = (\mathbf{s}-\hat{\mathbf{s}})^T(S^{-1}+N^{-1})(\mathbf{s}-\hat{\mathbf{s}}) + (\mathbf{q}-\hat{\mathbf{q}})^T C_q^{-1}(\mathbf{s}-\hat{\mathbf{q}}) + \mathbf{y}^T C_\perp^{-1}\mathbf{y} \tag{14}$$

where

$$\hat{\mathbf{s}} = S(S+N)^{-1}(\mathbf{y}-L\hat{\mathbf{q}}) \equiv SC^{-1}(\mathbf{y}-L\hat{\mathbf{q}}) \tag{15}$$

and

$$\hat{\mathbf{q}} = (L^TC^{-1}L)^{-1}L^TC^{-1}\mathbf{y} \equiv C_qL^TC^{-1}\mathbf{y} \tag{16}$$

are the best estimate for the intrinsic variability and the linear parameters, respectively, while

$$C_\perp^{-1} = C^{-1} - C^{-1}LC_qL^TC^{-1} \tag{17}$$

is the component of C that is orthogonal to the fitted linear functions. The variance in the estimate for the mean light curve is

$$\langle \mathbf{\Delta s}^2\rangle = S - S^TC_\perp S \tag{18}$$

and the variances for the linear parameters are

$$\langle \mathbf{\Delta q}^2\rangle = (L^TC^{-1}L)^{-1} \equiv C_q, \tag{19}$$

where $\mathbf{\Delta s} = \mathbf{s}-\hat{\mathbf{s}}$ and $\mathbf{\Delta q} = \mathbf{q}-\hat{\mathbf{q}}$, respectively.

We can marginalize the probability over the light curve $\mathbf{s}$ and the linear parameters $\mathbf{q}$ under the assumption of uniform priors for these variables to find that

$$P\left(\mathbf{y}|\mathbf{p}\right) \propto |S+N|^{-1/2}\left|L^TC^{-1}L\right|^{-1/2}\exp\left(-\frac{\mathbf{y}^TC_\perp^{-1}\mathbf{y}}{2}\right). \tag{20}$$

where the remaining parameters $\mathbf{p}$ are those describing the process (Eq. 1) and the transfer functions. The term $\mathbf{y}^TC_\perp^{-1}\mathbf{y}$ in the exponent is the generalized χ^2 that we calculated throughout the paper.

1.3 Fast Generation of Simulated Light Curves

We are interested in light curves constrained to resemble the observed continuum light curve. From Eq. (15) and Eq. (18), such a light curve is simply the estimated mean light curve given by Eq. (15) with a random component added that has the covariance matrix $Q = (S^{-1}+N^{-1})^{-1}$. [Rybicki & Press(1992)] suggest determining the eigenmodes of Q which are then the independent "normal" modes that can be added to the mean light curve to produce a random realization constrained by the continuum light curve. This is computationally expensive. Instead, we note that if we Cholesky decompose $Q = M^TM$, where M is an upper triangular matrix, and define the random component of the light curve by $\mathbf{u} = M\mathbf{r}$ where $\mathbf{r}$ is a vector of zero mean, unit dispersion Gaussian random deviates, that

$$\langle \mathbf{u}\mathbf{u}^T\rangle = M\langle \mathbf{r}\mathbf{r}^T\rangle M^T = MM^T = Q^T = Q \tag{21}$$

since the covariance matrix of the Gaussian deviates is simply the identity matrix and Q is symmetric. Since Q^{-1} is a tridiagonal matrix given the exponential covariance matrix and a diagonal noise matrix, we can generate very high dimension $\mathbf{u}$ that can be convolved with the transfer function to produce a simulated line light curve in $O(N)$ operations rather than the $O(N^3)$ needed following the eigenmode approach.

2 Transfer Function and Covariance Matrix Calculation

The expressions for the covariance matrices used in this paper and the accompanying code assuming transfer function of a simple top hat,

$$g(t-t') = A\,(t_2-t_1)^{-1} \quad \text{for} \quad t_1 \le t-t' \le t_2 \tag{22}$$

can be calculated analytically by substituting Eq. (22) and Eq. (1) into the integrals of Eq.(3,4 and 5), respectively.

2.1 Covariance Matrix of the Correlation Function Between Continuum and One Line

The covariance between continuum $S_c(t)$ at t_j and line $S_l(t)$ at t_i with transfer function defined as in Eq. (22) is

$$\langle s_c(t_j)s_l(t_i)\rangle = \tau\delta^2 A \begin{cases} e^{-\frac{t_L}{\tau}} - e^{-\frac{t_H}{\tau}} & \text{if } t_L > 0 \\ e^{\frac{t_H}{\tau}} - e^{\frac{t_L}{\tau}} & \text{if } t_H < 0 \\ 2 - e^{\frac{t_L}{\tau}} - e^{-\frac{t_H}{\tau}} & \text{if } t_L \le 0 \le t_H \end{cases} \tag{23}$$

where $t_L \equiv t_i - t_j - t_2$ and $t_H \equiv t_i - t_j - t_1$.

2.2 Covariance Matrix of the Autocorrelation of One Line

The covariance of line $S_l(t)$ between time t_i and t_j for the top hat transfer function is

$$\langle s_l(t_i)s_l(t_j)\rangle = \tau^2\delta^2 A^2 \begin{cases} (e^{\frac{-\Delta t}{\tau}}\,(e^{\frac{-t_M}{\tau}} + e^{\frac{t_M}{\tau}}) - 2\,e^{\frac{-|t_M|}{\tau}} + \begin{cases} 2\,\frac{\Delta t - t_M}{\tau} & \text{if } t_L \le 0 < t_M \\ 2\,\frac{\Delta t + t_M}{\tau} & \text{if } t_M \le 0 \le t_H \end{cases} \\ e^{\frac{-|t_M|}{\tau}}\,(e^{\frac{\Delta t}{2\tau}} - e^{\frac{-\Delta t}{2\tau}})^2 \qquad \text{if } t_L > 0 \text{ or } t_H < 0 \end{cases} \tag{24}$$

where

$$\begin{aligned} \Delta t &\equiv t_2 - t_1, \\ t_L &\equiv t_i - t_j - \Delta t, \\ t_M &\equiv t_i - t_j, \\ \text{and} \quad t_H &\equiv t_i - t_j + \Delta t. \end{aligned} \tag{25}$$

2.3 Covariance Matrix of the Correlation Function Between Two Lines

Consider the case when the first line $S_l(t)$ has transfer function $g(t-t')$ as defined in Eq. (22) and the other line $S_l'(t)$ has transfer function $g'(t-t')$

$$g'(t-t') = B\,(t_4-t_3)^{-1} \quad \text{for} \quad t_3 \le t-t' \le t_4 \tag{26}$$

where $t_4 - t_3 \le t_2 - t_1$, we have the covariance between line $S_l(t)$ at time t_i and line $S_l'(t)$ at time t_j

$$\langle s_l(t_i)s_l'(t_j)\rangle = \tau^2\delta^2 A\,B \begin{cases} e^{\frac{-|t_L|}{\tau}} + e^{\frac{-|t_H|}{\tau}} - e^{\frac{-|t_{M1}|}{\tau}} - e^{\frac{-|t_{M2}|}{\tau}} + \begin{cases} \frac{2\,t_H}{\tau} & \text{if } t_{M2} \le 0 < t_H \\ \frac{2\,(t_4-t_3)}{\tau} & \text{if } t_{M2} \le 0 < t_H \\ \frac{-2\,t_L}{\tau} & \text{if } t_L \le 0 < t_{M1} \end{cases} \\ e^{\frac{-|t_L|}{\tau}} + e^{\frac{-|t_H|}{\tau}} - e^{\frac{-|t_{M1}|}{\tau}} - e^{\frac{-|t_{M2}|}{\tau}} \qquad \text{if } t_L > 0 \text{ or } t_H < 0 \end{cases} \tag{27}$$

where

$$
\begin{aligned}
t_L &\equiv (t_i - t_j) - (t_2 - t_3), \\
t_{M1} &\equiv (t_i - t_j) - (t_2 - t_3), \\
t_{M2} &\equiv (t_i - t_j) - (t_1 - t_3), \\
\text{and} \quad t_H &\equiv (t_i - t_j) - (t_1 - t_4).
\end{aligned} \tag{28}
$$

References

[Rybicki & Press(1992)] Rybicki, G. B., & Press, W. H. 1992, The Astrophysical Journal, 398, 169