Lomb-Scargle Periodogram

EEL 6537

January, 2009

Outline

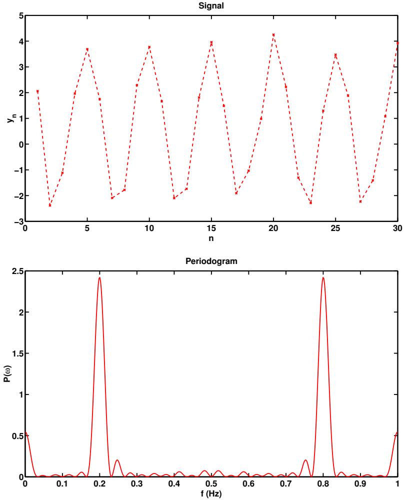
- Periodogram
- Lomb-Scargle Periodogram (P_x)
  - Brief introduction
  - Derivation from a least-squares fitting viewpoint
- Plain Least-Squares Periodogram (P_{LS})
  - Simpler than P_x
  - It has all nice properties that P_x has

Periodogram

- A diagram to identify the periodicity
- Provided samples $y_1, \ldots, y_N$, an estimate to the spectral density can be given by the periodogram:

$$P(\omega) = \frac{1}{N} \left| \sum_{n=1}^{N} y_n e^{-j\omega n} \right|^2$$

- $y_n = 3\cos(2\pi f n) + \eta_n$ $(n = 1, \ldots, 30)$ where $f = 0.2$ and $\eta \sim \mathcal{N}(0, 2)$.



Lomb-Scargle Periodogram P_x

- Identify real-valued sinusoids
- Lomb 1976 & Scargle 1982

$$P_x(\omega) = \frac{1}{2}\left\{ \frac{\left[\sum_{n=1}^{N} y(t_n)\cos(\omega(t_n-\tau))\right]^2}{\sum_{n=1}^{N}\cos^2(\omega(t_n-\tau))} + \frac{\left[\sum_{n=1}^{N} y(t_n)\sin(\omega(t_n-\tau))\right]^2}{\sum_{n=1}^{N}\sin^2(\omega(t_n-\tau))} \right\}$$

where τ (depending on ω) is defined as

$$\tan(2\omega\tau) = \frac{\sum_{n=1}^{N}\sin(2\omega t_n)}{\sum_{n=1}^{N}\cos(2\omega t_n)}$$

- coincides with the least-squares (LS) derivation
- has a clear statistical description
- has the invariance property to time translation

Derivation of $P_x(\omega)$ (Lomb 1976)

- Consider the following LS fitting problem:

$$\min_{\substack{\alpha \geq 0 \\ \bar{\phi} \in [0, 2\pi]}} \boxed{f(\alpha, \bar{\phi}) = \sum_{n=1}^{N} \left[y(t_n) - \alpha \cos(\omega t_n + \tilde{\phi} + \bar{\phi}) \right]^2}$$

where $\tilde{\phi}$ (over-parameterization) is subject to the constraint:

$$\sum_{n=1}^{N} \cos(\omega t_n + \tilde{\phi}) \sin(\omega t_n + \tilde{\phi}) = 0$$

$$\left(\text{equivalent to } \tan(2\tilde{\phi}) = -\frac{\sum_{n=1}^{N} \sin(2\omega t_n)}{\sum_{n=1}^{N} \cos(2\omega t_n)} \right)$$

- Define $a = \alpha\cos(\bar{\phi})$ and $b = -\alpha\sin(\bar{\phi})$:

$$f(\alpha, \bar{\phi}) = f(\boldsymbol{\theta}) = \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|$$

where

$$\mathbf{y} = \begin{bmatrix} y(t_1) \\ \vdots \\ y(t_n) \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} \cos(\omega t_1 + \tilde{\phi}) & \sin(\omega t_1 + \tilde{\phi}) \\ \vdots & \vdots \\ \cos(\omega t_n + \tilde{\phi}) & \sin(\omega t_n + \tilde{\phi}) \end{bmatrix}, \quad \boldsymbol{\theta} = \begin{bmatrix} a \\ b \end{bmatrix}$$

- $\min f(\boldsymbol{\theta}) \Rightarrow \hat{\boldsymbol{\theta}} = (\mathbf{A}^T\mathbf{A})^{-1}(\mathbf{A}^T\mathbf{y})$

$$\boxed{P_x(\omega) = \frac{1}{N}(\mathbf{A}\hat{\boldsymbol{\theta}})^T(\mathbf{A}\hat{\boldsymbol{\theta}})}$$

- By using the equality related to $\tilde{\phi}$, we obtain the Lomb-Scargle Periodogram (up to a multiplicative factor)

$$P_x(\omega) = \frac{1}{N}\left\{ \frac{\left[\sum_{n=1}^{N} y(t_n)\cos(\omega t_n + \tilde{\phi})\right]^2}{\sum_{n=1}^{N}\cos^2(\omega t_n + \tilde{\phi})} + \frac{\left[\sum_{n=1}^{N} y(t_n)\sin(\omega t_n + \tilde{\phi})\right]^2}{\sum_{n=1}^{N}\sin^2(\omega t_n + \tilde{\phi})} \right\}$$

where $\tilde{\phi}$ is subject to

$$\tan(2\tilde{\phi}) = -\frac{\sum_{n=1}^{N}\sin(2\omega t_n)}{\sum_{n=1}^{N}\cos(2\omega t_n)} \qquad \blacksquare$$

The Statistical Property of P_x

- If $\{y(t_n)\}$ i.i.d. $\sim \mathcal{N}(0,1)$, then $NP_x \sim \exp(2)$
- Proof:

$$\sum_{n=1}^{N} y(t_n)\cos(\omega t_n + \tilde{\phi}) \sim \mathcal{N}\left(0, \sum_{n=1}^{N} \cos^2(\omega t_n + \tilde{\phi})\right)$$

$$\Longrightarrow C = \frac{\sum_{n=1}^{N} y(t_n)\cos(\omega t_n + \tilde{\phi})}{\sqrt{\sum_{n=1}^{N} \cos^2(\omega t_n + \tilde{\phi})}} \sim \mathcal{N}(0,1)$$

$$\text{similarly } S = \frac{\sum_{n=1}^{N} y(t_n)\sin(\omega t_n + \tilde{\phi})}{\sqrt{\sum_{n=1}^{N} \sin^2(\omega t_n + \tilde{\phi})}} \sim \mathcal{N}(0,1)$$

$$NP_x = C^2 + S^2 \sim \chi^2(2) \sim \exp(2) \qquad \blacksquare$$

An Alternative to Lomb-Scargle Periodogram

- P_x is favored: [Lomb 1975] cited 994 times & [Scargle 1982] cited 1784 times
- How about plain least-squares periodogram without the $\tilde{\phi}$ over-parameterization?

$$\min_{\substack{\alpha \geq 0 \\ \phi \in [0, 2\pi]}} \boxed{f(\alpha, \phi) = \sum_{n=1}^{N} [y(t_n) - \alpha \cos(\omega t_n + \phi)]^2}$$

Plain Least-Squares Periodogram P_{LS}

- Define

$$\mathbf{y} = \begin{bmatrix} y(t_1) \\ \vdots \\ y(t_n) \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} \cos\omega t_1 & \sin\omega t_1 \\ \vdots & \vdots \\ \cos\omega t_n & \sin\omega t_n \end{bmatrix}, \quad \boldsymbol{\theta} = \begin{bmatrix} \alpha\cos(\phi) \\ -\alpha\sin(\phi) \end{bmatrix}$$

we obtain

$$\min_{\boldsymbol{\theta}} f(\boldsymbol{\theta}) = \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|^2 \Rightarrow \hat{\boldsymbol{\theta}} = (\mathbf{A}^T\mathbf{A})^{-1}(\mathbf{A}^T\mathbf{y})$$

- Plain least-squares periodogram

$$\boxed{P_{\text{LS}}(\omega) = \frac{1}{N}(\mathbf{A}^T\hat{\boldsymbol{\theta}})^T(\mathbf{A}^T\hat{\boldsymbol{\theta}}) = \frac{1}{N}\mathbf{r}^T\mathbf{R}^{-1}\mathbf{r}}$$

where $\mathbf{R} = \mathbf{A}^T\mathbf{A}$ and $\mathbf{r} = \mathbf{A}^T\mathbf{y}$

UF UNIVERSITY of FLORIDA

The Statistical Property of P_{LS}

- If $\{y(t_n)\}$ i.i.d. $\sim \mathcal{N}(0,1)$, then $NP_{\text{LS}} \sim \exp(2)$

- Proof: (recall that $\mathbf{R} = \mathbf{A}^T\mathbf{A}$ and $\mathbf{r} = \mathbf{A}^T\mathbf{y}$)

$$\mathbf{y} \sim \mathcal{N}(0, \mathbf{I}) \;\Rightarrow\; \mathbf{r} \sim \mathcal{N}(0, \mathbf{R})$$

$$\Rightarrow \mathbf{R}^{-1/2}\mathbf{r} \sim \mathcal{N}(0, \mathbf{I})$$

$$\Rightarrow NP_{\text{LS}} = \mathbf{r}^T\mathbf{R}^{-1}\mathbf{r} = \|\mathbf{R}^{-1/2}\mathbf{r}\|^2 \sim \chi^2(2)$$

Time Invariance Property of P_{LS}

- Rewrite the LS fitting problem

$$\min_{\substack{\alpha \geq 0 \\ \phi \in [0, 2\pi]}} f(\alpha, \phi) = \sum_{n=1}^{N} [y(t_n) - \alpha \cos(\omega t_n + \phi)]^2$$

as

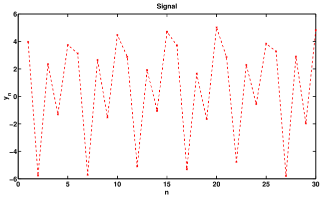
$$\min_{\beta(\omega)} \boxed{\sum_{n=1}^{N} \left| y(t_n) - \left(\beta e^{j\omega t_n} + \beta^* e^{-j\omega t_n}\right) \right|^2}$$

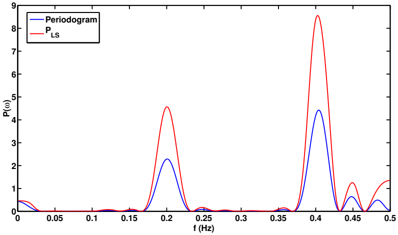
where $\beta = |\beta| e^{j\phi}$ and $2|\beta| = \alpha$

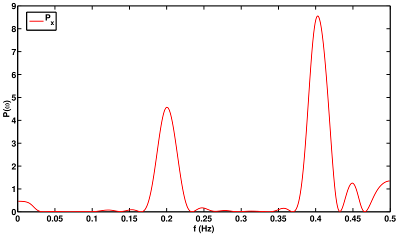
- A time shift Δt (i.e. t replaced by $t + \Delta t$) will only introduce a phase shift $e^{j\omega \Delta t}$ to $\beta(\omega)$. P_{LS} is not affected.

P_x and P_{LS} give the same estimate!

- $y_n = 3\cos(2\pi f_1 n + \phi_1) + 4\cos(2\pi f_2 n + \phi_2) + \eta_n$ $(n = 1, \ldots, 30)$ where $f_1 = 0.2$, $f_2 = 0.4$, ϕ_1 and ϕ_2 are randomly chosen and $\eta \sim \mathcal{N}(0, 2)$







Conclusion

- The Lomb-Scargle Periodogram P_x and the Plain Least-Squares Periodogram P_{LS} give the same estimate, because the best least-squares fitting is unique.
- P_x has been favored so far for unclear grounds
- P_{LS} is simpler than P_x but with the same nice properties
- We recommend using P_{LS}

UF UNIVERSITY of FLORIDA