Vorlesung “Astronomische Datenanalyse”

Praktikum 3: Zeitreihenanalyse

Version vom 11. Juli 2002

Das Praktikum findet im Institut für Astronomie und Astrophysik, Abt. Astronomie im Gebäude Sand 1 statt. Unter der URL `http://astro.uni-tuebingen.de/~wilms/teach/data/` finden Sie PS- und PDF-Dateien dieser Anleitung, die Sie sich ausdrucken können, sowie alle hier angegebenen Programme.

1 Zeitreihenanalyse

Auch wenn der Himmel vielen Menschen als statisch erscheint, sind zeitlich veränderliche Phänomene in der Astronomie eher die Regel als die Ausnahme. Sie können in fast allen astronomischen Objekten beobachtet werden, sei es in variablen Sternen, in Radiopulsaren, in akkretierenden Neutronensternen im Röntgenbereich oder in den Lichtkurven aktiver galaktischer Kerne.

In diesem Praktikum werden wir uns mit dem zur Charakterisierung von Zeitreihen am häufigsten verwendeten Werkzeug, dem Periodogramm, beschäftigen. Alle benötigten Daten finden Sie unter `http://astro.uni-tuebingen.de/~wilms/teach/data/prakt3/`.

1.1 Definitionen

Zeitreihenanalyse wird häufig mit den Techniken der Fourier-Analyse betrieben. Hier werden wir uns auf die Zeitreihenanalyse einer äquidistanten Zeitreihe beschränken, also einer Zeitreihe, bei der die zeitliche Information x_i an den Zeitpunkten $t_i = t_0 + i\Delta t$ gemessen wurde. Hier ist i eine natürliche Zahl und t_0 ist der Nullpunkt der Zeitreihe. In der Fourieranalyse wird gezeigt, daß jede Zeitreihe dargestellt werden kann als

$$x_k = \frac{1}{N} \sum_{k=0}^{N-1} X_j \mathrm{e}^{-2\pi i jk/N} \tag{1}$$

wo

$$X_j = \sum_{k=0}^{N-1} x_k \mathrm{e}^{2\pi i jk/N} \tag{2}$$

Die Amplituden X_j bezeichnet man als die (diskrete) Fourier-Transformation von x_i. Die Frequenz, für die X_j berechnet wird,

$$\omega_j = 2\pi\nu_j = 2\pi j/(N\Delta t) \tag{3}$$

ist die Fourier-Frequenz. Die höchste Frequenz, für die X_j berechnet werden kann, $\nu_{N/2} = 1/(2\Delta t)$, nennt man die Nyquist-Frequenz. Der zur Frequenz 0 gehörige Term ist proportional zum Mittelwert der Zeitreihe. Häufig wird dieser Fourierkoeffizient explizit auf 0 gesetzt, indem man von der Zeitreihe vor der Fourier-Transformation ihren Mittelwert abzieht.

In der praktischen Analyse von Zeitreihen ist es häufig nicht notwendig, die ganze in der trigonometrischen Summe (1) steckende Information zu berücksichtigen. Für viele Anwendungen langt es, die Amplitude der an einer bestimmten Frequenz zu betrachten, die gegeben ist durch

$$P_j = A|X_j|^2 \quad \text{wo } j \in [0 \ldots N/2] \tag{4}$$

Diese Werte werden als Periodogramm bezeichnet. Weitere gebräuchliche Namen sind PSD für *power spectral density* oder PDS für *power density spectrum*. Da astronomische Messungen normalerweise reelle Zahlen sind, gilt $X_j = X^*_{-j}$ und daher $|X_j|^2 = |X_{-j}|^2$, so daß das Periodogramm nur für positive Frequenzen definiert werden muß (der Stern, *, bezeichnet die komplexe Konjungation).

Unglücklicherweise ist die Normierungskonstante A des Periodogramms nicht standardisiert. Wir werden hier mit der sogenannten Leahy-Normierung arbeiten. In der Leahy-Normierung ist

$$A = \frac{2\Delta t}{X_0} = \frac{2\Delta t^2}{N_{\mathrm{ph}}} \tag{5}$$

Im zweiten Ausdruck ist N_{ph} die Gesamtzahl der detektierten Photonen, dieser Ausdruck gilt, wenn für x_i die gemessene Zählrate zum Zeitpunkt t_i ist. Die Leahy-Normierung wird in vielen Bereichen deshalb verwendet, weil der Wert des Periodogramms bei einer Zeitreihe, deren Variabilität ausschließlich durch die Poisson-Statistik charakteristiert ist, gleich 2 ist. Daher eignet sich das Periodogramm in Leahy-Normierung gut zur Bestimmung von Abweichungen von der Poisson-Statistik, d.h. sie eignet sich dazu, das der Quelle inhärente Signal von dem durch den Meßprozess induziertem statistischen Rauschen zu trennen.

1.2 Berechnung des Periodogramms

In IDL können Sie das Periodogramm in verschiedenen Normierungen mit der in Tübingen entwickelten Routine `PSD` bestimmen. Der Aufruf von `PSD` ist

```
psd,time,rate,freq,psd,/leahy
```

wo `time` die Zeit und `rate` die gemessene Zählrate ist. In den Arrays `freq` und `psd` wird die Fourier-Frequenz und der Wert des Periodogramms zurückgeliefert. Das Keyword `/leahy` stellt die korrekte Normierung des Periodogramms sicher. Außer diesen Argumenten hat die Routine noch weitere Argumente, die wir hier nicht betrachten werden.

Um die Eigenschaften des PSD kennenzulernen, werden wir in IDL zunächst einfache Zeitreihen generieren und uns deren Periodogramme betrachten. In einem zweiten Schritt werden wir dann die Zeitreihen zweier typischer astronomischer Objekte analysieren.

1.3 Einfache Periodogramme

Der einfachste Fall einer Zeitreihe ist eine rein sinusförmige Variabilität. Diese können wir in IDL mit dem Programm `zeitsinus` untersuchen:

```
;;
;; Periodogramm einer sinusfoermigen Zeitreihe
;;

tmin=0  ;; Anfangszeit
tmax=100.;; Endzeit
f=2.    ;; Frequenz
r0=1    ;; Mittlere Rate
ampl=2. ;; Amplitude

npt=600 ;; Zahl der Punkte in der Zeitreihe
time=tmin+(tmax-tmin)*findgen(npt)/(npt-1)
rate=r0+ampl*sin(2.*!pi*f*time)
psd,time,rate,fre,psd,/leahy
plot,fre,psd

END
```

Ändern Sie `npt` zu kleineren und größeren Werten und beobachten Sie, wie sich der Fourier-Peak ändert. Modifizieren Sie dann das Programm `zeitsinus` und berechnen Sie die Periodogramme komplizierterer Zeitreihen:

Mehrere Sinusse:

$$x(t) = r_0 + A_1 \sin(2\pi f_1 t) + A_2 \sin(2\pi f_2 t)$$

Was passiert, wenn Sie A_1 und A_2 variieren? Welche Frequenzen f_1 und f_2 können Sie im Periodogramm noch trennen?

Zeitreihe mit exponentiell abfallender Amplitude:

$$x(t) = r_0 + A_1 \exp(-0.69\, t/t_{1/2}) \sin(2\pi f_1 t)$$

Betrachten sie hier das PSD bei verschiedenen Halbwertszeiten $t_{1/2}$ und variieren Sie das Verhältnis von f_1 und $t_{1/2}$.

Solche periodischen Funktionen sind allerdings eher selten. In der Astronomie ist der Meßprozeß häufig durch starkes Rauschen beeinträchtigt. Sie können den Effekt dieses Rauschens in Ihren Simulationen durch die Addition von normalverteilten Zufallszahlen zur simulierten Lichtkurve studieren:

```
seed=13244L ;; variieren!
rate=rate+a1*randomu(seed,n_elements(rate),/normal)
```

wo `a1` die Standardabweichung des Rauschens ist. Wiederholen Sie die obigen Experimente unter Berücksichtigung dieser Rauschterme! Bis zu welcher Stärke des Rauschens können Sie die Periode noch in den simulierten Daten wiederfinden? Betrachten Sie auch die Lichtkurve – können Sie die Periode, die Sie in die Daten gesteckt haben, noch mit bloßem Auge wiederfinden?

1.4 Astronomische Periodogramme

1.4.1 4U 0115+62

Im Frühjahr 1999 brach der Be-Röntgendoppelstern 4U 0115+62 erneut aus. In solchen Systemen bewegt sich ein Neutronenstern auf einer exzentrischen Bahn um einen sehr schnell rotierenden B-Stern. Die rotation des B-Sterns is so schnell, daß am Äquator des Sterns fast die Fluchtgeschwindigkeit erreicht wird. Dadurch bildet sich um den Stern eine Gasscheibe aus, die im optischen Spektrum durch Emissionslinien detektierbar ist (daher das “e” im Namen des Systems). Beim Durchgang des Neutronensterns durch eine solche Gasscheibe kann Material auf den Neutronenstern akkretieren und den beobachteten Röntgenausbruch erzeugen. Warum dies nicht bei jeder Periastronpassage passiert ist nicht geklärt.

In Zusammenarbeit mit Kollegen aus San Diego haben wir den Ausbruch von 1999 mit dem Röntgensatelliten RXTE täglich beobachtet. Einzelheiten dieser Beobachtungen finden Sie unter `http://astro.uni-tuebingen.de/publications/ppC_99.shtml`. Sie sollten dieses Paper auf dem Astrophysical Data System (`http://adsabs.harvard.edu/default_service.html`) suchen und in seiner veröffentlichten Form ausdrucken. Dort finden Sie weitere Hinweise, nach was Sie im folgenden zu suchen haben ...

In der Datei `src-0.125s-2nd-long.lc` finden Sie die Lichtkurve von 4U 0115+62 vom 12. März 1999. Diese Lichtkurve hat eine Auflösung von 0.125 s und wurde während zweier Umläufe von RXTE um die Erde aufgenommen.

Mit der IDL-Unterroutine `readlc` können Sie diese Daten in IDL einlesen:

```
readlc,time,counts,'src-0.125s-2nd-long.lc',/mjd,/counts
```

wo `time` die Meßzeit ist (hier in *Tagen* angegeben).

Konvertieren Sie zunächst die Zeit in Sekunden seit dem Beginn der Messung:

```
time=(time-time[0])*86400.
```

Sie sehen, daß Sie nichts sehen ... Die Lichtkurve ist zeitlich zu gut aufgelöst. Verwenden Sie daher den `xrange`-Parameter des Plot-Befehls, um sich einen Teil der Lichtkurve zu betrachten. Gute Beobachtungslängen sind Stücke von 1000 s Länge oder noch wesentlich kürzer. Welche Arten zeitlicher Phänomene können Sie mit dem Auge detektieren?

Die gemessenen Daten enthalten eine Lücke, die durch die Zeit, während derer der Neutronenstern vom Satelliten aus gesehen hinter der Erde war. Wir wollen im folgenden nur das längere Zeitsegment nach der Lücke betrachten. Dazu müssen wir diesen Teil der Zeitreihe aus der Lichtkurve herausschneiden. Bestimmen Sie dazu die Länge der einzelnen Zeitbins:

```
dt=shift(time,-1)-time
```

Die Lücke ist logischerweise dort, wo `dt` wesentlich größer ist als die Binzeit der Zeitreihe. Bestimmen Sie die Lage der Lücke und schneiden Sie den langen Teil der Zeitreihe aus der Lichtkurve aus. Dazu müssen Sie beachten, daß aufgrund von Rundungsfehlern einige Zeitintervalle eine unwesentlich andere Länge haben. Da die Lücke das längste Intervall ist, ist es am einfachsten, sie mit IDLs `max`-Prozedur zu finden:

```
dummy=max(dt,ndx)
ndx=ndx+1
time=time[ndx:n_elements(time)-1]
counts=counts[ndx:n_elements(counts)-1]
dt=dt[ndx:n_elements(dt)-1]
```

Schlußendlich ist es noch notwendig, anstelle der Gesamtzahl der in einem Zeitintervall gemessenen Photonen die Zählrate zu verwenden:

```
rate=counts/dt
```

Berechnen Sie jetzt von dieser Lichtkurve das Periodogramm! Es empfiehlt sich übrigens, solche Periodogramme doppeltlogarithmisch darzustellen, was Sie mit dem `/xlog` und dem `/ylog` Keyword der `plot`-Routine erreichen können.

Ein Nachteil des eben berechneten Periodogramms ist es, daß es einen sehr großen Frequenzbereich abdeckt und sehr verrauscht ist. Wie gezeigt werden kann ist der statistische Fehler des Periodogramms so groß wie sein Wert. Daher ist es häufig nicht sinnvoll, Einzelperiodogramme zu betrachten. Vielmehr werden viele Periodogramme von Stücken der Lichtkurve berechnet, die dann gemittelt werden. Daber wird der statistische Fehler des Periodogramms vermindert, auch wenn gleichzeitig der Frequenzbereich eingeschränkt wird. Die Länge dieser Segmente (in Einheiten von `dt`) können Sie mit dem Keyword `dseg` zur `psd`-Unterroutine setzen. Berechnen Sie daher PSDs der Lichtkurve für unterschiedliche Werte von `dseg`. Es empfiehlt sich aus Rechenzeitgründen, für `dseg` Potenzen von 2 zu wählen, da der zugrundeliegende FFT-Algorithmus bei diesen *sehr* viel schneller arbeitet ($O(N \log N)$), als bei zufällig gewählten Längen ($O(N^2)$).

Da Sie bei verschiedenen Längen der Lichtkurve unterschiedlich breite Frequenzbereiche haben ist es sinnvoll, das endgültige Periodogramm aus verschiedenen Periodogrammen zusammenzusetzen. Erweitern Sie ihr Programm dementsprechend. Welche Frequenzen können Sie feststellen?

Eine der Frequenzen, die Sie gerade bestimmt haben, ist die Rotationsfrequenz des Neutronensterns. Bestimmen Sie diese Frequenz aus dem Periodogramm und berechnen Sie das Pulsprofil. Dazu müssen Sie für jedes Zeitbin die zugehörige Pulsphase bestimmen (siehe letztes Praktikum). Definieren Sie sich gleichabständige Pulsphasen und berechnen Sie die mittlere Zählrate für jedes Phasenintervall um das Pulsprofil zu erhalten.

1.5 Die Sonne

Wie ja in der Vorlesung besprochen wurde, ist die Variation der Sonnenfleckenrelativzahl ein besonders schönes Beispiel für einen periodischen Vorgang in der Astronomie. Die in der Vorlesung gezeigten Daten können Sie in der Datei `spot_num.txt` finden – lesen Sie die Daten ein und berechnen Sie das Periodogramm!

1.6 Ein variabler Stern

Wenn Ihnen die Sonne zu nahe ist, dann können Sie alternativ zur Bestimmung der Sonnenfleckenperiode auch den variablen Stern HL Tau bearbeiten, dessen Lichtkurve in der Datei `HLTau_all` zu finden ist.