Non-Imaging Data Analysis

Greg Taylor

Based on the original lecture by T.J. Pearson

Introduction

• Reasons for analyzing visibility data
  • Insufficient (u,v) plane coverage to make an image
  • Inadequate calibration
  • Quantitative analysis
  • Direct comparison of two data sets
  • Error estimation
    - Usually, visibility measurements are independent gaussian variates
    - Systematic errors are usually localized in the (u,v) plane
  • Statistical estimation of source parameters

Inspecting Visibility Data

• Fourier imaging
  \[ V(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(l, m) f(l, m) \exp[-2\pi i(lu + mv)] \, dl \, dm \]

• Problems with direct inversion
  Sampling
    - Poor (u,v) coverage
    - Missing data
    - e.g., no phases (speckle imaging)
  Calibration
    - Closure quantities are independent of calibration
  Non-Fourier imaging
    - e.g., wide-field imaging; time-variable sources (SS433)
    - Noise
    - Noise is uncorrelated in the (u,v) plane but correlated in the image

Useful displays
  • Sampling of the (u,v) plane
  • Amplitude and phase vs. radius in the (u,v) plane
  • Amplitude and phase vs. time on each baseline
  • Amplitude variation across the (u,v) plane
  • Projection onto a particular orientation in the (u,v) plane

Example: 2021+614
  • GHz peaked spectrum radio galaxy at z=0.23
  • A VLBI dataset with 11 antennas from 1987
  • VLBA only in 2000

- Fourier Transform theorems
  - Linearity
  - Viabilities of components add (complex)
  - Convolution
  - Shift
    - Shifting the source creates a phase gradient across the (u,v) plane
  - Similarity
    - Larger sources have more compact transforms

\[ F(\omega, \theta) = \int_{-\infty}^{\infty} f(x) e^{i \omega x} e^{i \theta y} \, dx \, dy \]

**Properties of the Fourier transform**

- Linearity
- Viabilities of components add (complex)
- Convolution
- Shift
- Shifting the source creates a phase gradient across the (u,v) plane
- Similarity
- Larger sources have more compact transforms

**Fourier Transform theorems**

\[ F(\omega, \theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{i \omega x} e^{i \theta y} \, dx \, dy \]

- Linearity
  - \( F(a f(x)) = \frac{1}{|a|} F(f(x)) \)
- Convolution
  - \( F(f(x) * g(x)) = F(f(x)) * F(g(x)) \)
- Shift
  - \( F(f(x - a)) = F(f(x)) e^{-i \omega a} \)
Visibility function

Brightness distribution

Visibility at short baselines contains little information about the profile of the source.

Simple models

By inspection, we can derive a simple model:
Two equal components, each 1.25 Jy, separated by about 6.8 milliarcsec in p.a. 33º, each about 0.8 milliarcsec in diameter (gaussian FWHM)

To be refined later...

Projection in the (u,v) plane
**Closure Phase and Amplitude: closure quantities**

- Antenna-based gain errors
  \[ V_{\text{local}} = |V_{\text{local}}| \exp(i\phi_{\text{local}}) = \hat{G} \exp(i\phi_{\text{true}}) \exp(-i\phi_{\text{antenna}}) \]

- **Closure phase (bispectrum phase)**
  \[ \phi_{\text{closure}}(t) = \phi_{\text{true}}(t) + \phi_{\text{error}}(t) \]

- **Closure amplitude**
  \[ \frac{|V_{\text{closure}}|}{V_{\text{local}} / V_{\text{true}}} \]

  - Closure phase and closure amplitude are unaffected by antenna gain errors
  - They are conserved during self-calibration
  - Contain (N-2)/N of phase, (N-3)/(N-1) of amplitude info
    - Many non-independent quantities
    - They do not have gaussian errors
    - No position or flux info

**Model fitting**

- Imaging as an Inverse Problem
  - In synthesis imaging, we can solve the forward problem: given a sky brightness distribution, and knowing the characteristics of the instrument, we can predict the measurements (visibilities), within the limitations imposed by the noise.
  - The Inverse problem is much harder, given limited data and noise: the solution is rarely unique.
  - A general approach to Inverse problems is model fitting. See, e.g., Press et al., Numerical Recipes
    1. Design a model defined by a number of adjustable parameters.
    2. Solve the forward problem to predict the measurements.
    3. Choose a figure-of-merit function, e.g., rms deviation between model predictions and measurements.
    4. Adjust the parameters to minimize the merit function.

  - **Goals:**
    1. Best-fit values for the parameters.
    2. A measure of the goodness-of-fit of the optimized model.

**Uses of model fitting**

- Model fitting is most useful when the brightness distribution is simple.
  - Checking amplitude calibration
  - Starting point for self-calibration
  - Estimating parameters of the model (with error estimates)
  - In conjunction with CLEAN or MEM
  - In astrometry and geodesy

- Programs
  - AIPS UVFITY
  - Dfmap (Martin Shepherd)

**Maximum Likelihood and Least Squares**

- The model:
  \[ V(x, y) = F(x_0, y_0, \ldots, \sigma_M) + \text{noise} \]

- The likelihood of the model (if noise is gaussian):
  \[ L \propto \prod_{i=1}^{N} \exp \left\{ -\frac{1}{2} \left( \frac{V_{\text{true}} - F(x_0, y_0, \ldots, \sigma_M)}{\sigma_i} \right)^2 \right\} \]

- Maximizing the likelihood is equivalent to minimizing chi-square (for gaussian errors):
  \[ \chi^2 = \sum_{i=1}^{N} \left( \frac{V_{\text{true}} - F(x_0, y_0, \ldots, \sigma_M)}{\sigma_i} \right)^2 \]

- Follows chi-square distribution with N - M degrees of freedom. Reduced chi-square has expected value 1.

**Parameters**

- **Example**
  - Component position: (x,y) or polar coordinates
  - Flux density
  - Angular size (e.g., FWHM)
  - Axial ratio and orientation (position angle)
  - For a non-circular component
    - 6 parameters per component, plus a “shape”

This is a conventional choice: other choices of parameters may be better!

(Wavelets; shapelets? [Hermite functions])

Practical model fitting: 2021

Model fitting 2021

Limitations of least squares

Least-squares algorithms
Problems with least squares

- Global versus local minimum
- Slow convergence: poorly constrained model
  Do not allow poorly-constrained parameters to vary
- Constraints and prior information
  Boundaries in parameter space
  Transformation of variables
- Choosing the right number of parameters: does adding a parameter significantly improve the fit?
  Likelihood ratio or F test: use caution

Monte Carlo methods

Error estimation

- Find a region of the M-dimensional parameter space around the best fit point in which there is, say, a 68% or 95% chance that the true parameter values lie.
- Constant chi-square boundary: select the region in which
  \( \chi^2 < \chi^2_{\text{min}} + \Delta \chi^2 \)
- The appropriate contour depends on the required confidence level and the number of parameters estimated.
- Monte Carlo methods (simulated or mock data): relatively easy with fast computers
- Some parameters are strongly correlated, e.g., flux density and size of a gaussian component with limited (uv) coverage.
- Confidence intervals for a single parameter must take into account variations in the other parameters ("marginalization")

Problem: to detect changes in component positions between observations and measure their speeds
- Direct comparison of images is bad different (uv) coverage, uncertain calibration, insufficient resolution
- Visibility analysis is a good method of detecting and measuring changes in a source; allows "controlled super-resolution"
- Calibration uncertainty can be avoided by looking at the closure quantities: have they changed?
- Problem of differing (uv) coverage: compare the same (uv) points whenever possible
- Model fitting as an interpolation method

Applications: Superluminal motion

- Example 1: Discovery of superluminal motion in 3C279 (Whitney et al., Science, 1971)
  - 1.55 ± 0.03 milliarcsec in 4 months: \( \nu/c = 10 \pm 3 \)

Superluminal motion

- Example 2: Discovery of superluminal motion in 3C279 (Whitney et al., Science, 1971)
  - 1.55 ± 0.03 milliarcsec in 4 months: \( \nu/c = 10 \pm 3 \)

Example 3: Discovery of superluminal motion in 3C279 (Whitney et al., Science, 1971)

1.55 ± 0.03 milliarcsec in 4 months: \( \nu/c = 10 \pm 3 \)
Applications: Expanding sources

- Example 2: changes in the radio galaxy 2021+614 between 1987 and 2000
  - We find a change of 200 microarcsec so \( v/c = 0.18 \)
  - By careful combination of model-fitting and self-calibration, Conway et al. (1994) determined that the separation had changed by 69 \( \pm \) 10 microarcsec between 1982 and 1987, for \( v/c = 0.19 \)

GRB 030329

June 20, 2003

- \( t+83 \) days
- Peak \( \sim 3 \) mJy
- Size \( 0.172 \pm 0.043 \) mas
- \( 0.5 \pm 0.1 \) pc
- Average velocity \( = 3c \)

Taylor et al. 2004

VLBA+Y27+GBT+EB+AR+WB = 0.11 km^2

Proper motion limits

RA = \( -0.02 \pm 0.00 \) mas/yr
DEC = \( -0.44 \pm 0.03 \) mas/yr

Motion < 0.28 mas in 80 days

GRB 030329 subtracted
Applications: Gravitational Lenses

- Gravitational Lenses
  - Single source, multiple images formed by intervening galaxy.
  - Can be used to map mass distribution in lens.
  - Can be used to measure distance of lens and \( H_0 \); need redshift of lens and background source, model of mass distribution, and a time delay.

- Application of model fitting
  - Lens monitoring to measure flux densities of components as a function of time.
  - Small number of components, usually point sources.
  - Need error estimates.

  - VLA configuration changes: different HA on each day
  - Other sources in the field

1608 monitoring results

\[
\begin{align*}
B - A &= 31 \text{ days} \\
B - C &= 36 \text{ days} \\
H_0 &= 59 \pm 8 \text{ km/s/Mpc}
\end{align*}
\]

- The Sunyaev-Zeldovich effect
  - Photons of the CMB are scattered to higher frequencies by hot electrons in galaxy clusters, causing a negative brightness decrement.
  - Decrement is proportional to integral of electron pressure through the cluster, or electron density if cluster is isothermal.
  - Electron density and temperature can be estimated from X-ray observations, so the linear scale of the cluster is determined.
  - This can be used to measure the cluster distance and \( H_0 \).

- Application of model fitting
  - The profile of the decrement can be estimated from X-ray observations (beta model).
  - The Fourier transform of this profile increases exponentially as the interferometer baseline decreases.
  - The central decrement in a synthesis image is thus highly dependent on the (u,v) coverage.
  - Model fitting is the best way to estimate the true central decrement.

SZ profiles

SZ images

Reese et al. astro-ph/0205350
Summary

- For simple sources observed with high SNR, much can be learned about the source (and observational errors) by inspection of the visibilities.
- Even if the data cannot be calibrated, the closure quantities are good observables, but they can be difficult to interpret.
- Quantitative data analysis is best regarded as an exercise in statistical inference, for which the maximum likelihood method is a general approach.
- For gaussian errors, the ML method is the method of least squares.
- Visibility data (usually) have uncorrelated gaussian errors, so analysis is most straightforward in the (u,v) plane.
- Consider visibility analysis when you want a quantitative answer (with error estimates) to a simple question about a source.
- Visibility analysis is inappropriate for large problems (many data points, many parameters, correlated errors); standard imaging methods can be much faster.