How to Reconstruct the Original Shape of a Radar Signal?

Matthew G. Averill, Gang Xiang, Vladik Kreinovich, G. Randy Keller, and Scott A. Starks
Pan-American Center for Earth and Environmental Studies
University of Texas at El Paso, El Paso, TX 79968, USA
averill@geo.utep.edu, gxiang@utep.edu, vladik@utep.edu, keller@utep.edu, sstarks@utep.edu

Patrick S. Debroux and James Boehm
Army Research Laboratory, SLAD

1. Radars are Important

- Radar measurements are used in many areas of science and engineering.
- Historically the first use of radars was in tracing *airplanes* and *missiles*.
- This is still one of the main uses of radars.
- However, radars are used more and more in *geosciences* as well.
- The information provided by airborne radars nicely supplements other remote sensing information
  - radar beams can go below the leaves, to the actual Earth surface
  - and they can even go even deeper than the surface.

2. Main Use of Radars: Localization

- The main idea behind a radar is simple:
  - we send a pulse-like radio signal,
  - this signal gets reflected by the target, and
  - we measure the reflected signal.
- The main information that we can get from the radar is the *travel time*.
- Based on the travel time, we can find the distance between the radar and the target.
- If we use several radars, we can thus get an exact location of the target.
- This is how radars determine the exact position of the planes in the vicinity of an airport.
- This is how radars produce high-accuracy digital elevation maps that are so important in geophysics.

3. Radars Provide Additional Information

- If the targets were *points*, then after sending a pulse signal, we would get a pulse back.
- In this case, the only information we are able to get is the *distance* from the radar to the point target.
- In reality, the target is *not* a point.
- As a result:
  - even if we send a pulse signal,
  - this pulse is reflected from different points on a target and
  - therefore, we get a continuous signal back.
- The shape of this signal can provide us with the additional information about the reflecting surface.

4. It Is Desirable to Determine the Probability Distribution of the Reflected Signal

- In an airborne geophysical radar, pulses are sent one after another.
- As a result, individual reflections get entangled.
- We can still measure the *probability distribution* of the values of the reflected signal.
- Our *objective*: to extract the information about the reflecting surface from this distribution.

5. Filtering

- *Problem:* the reflected signals are weak and covered with noise.
- *Solution:* to decrease the noise, we apply *filtering* – usually, linear filtering.
- *What is linear filtering:*
  - instead of the original signal $x(t)$,
  - we consider a linear combination of this signal and the signals ate the previous moments of time:

$$y(t) = \sum_s a(s) \cdot x(t-s).$$

+ This filtering decreases the noise and makes the distance measurement very accurate.

– On the other hand, it replaces the original possibly non-Gaussian signal $x(t)$ with a linear combination of such signals.

6. Filtering: A Problem

- *Central limit theorem:* as we increase the number of terms in a linear combination of several small random variables, the resulting distribution of a sum tends to Gaussian.
- *Comment:* this theorem is the main reason why Gaussian distributions are so frequent in practice.
- *Conclusion:* after filtering, we get a distribution that is close to Gaussian.
- *Problem:*
  - we have a probability distribution for
  $$y(t) = \sum_s a(s) \cdot x(t-s);$$
  - we want to reconstruct the original distribution for x.

7. Main Idea: Use Logarithmic Moments

- *Main idea:* describe both distribution in terms of logarithmic moments.
- *Characteristic function* of a random variable ξ is

$$\chi_\xi(\omega) \stackrel{\text{def}}{=} E[\exp(\mathrm{i}\cdot\omega\cdot\xi)].$$

- For the sum $\xi = \xi_1 + \xi_2$ of independent ξ_i,

$$E[\exp(\mathrm{i}\cdot\omega\cdot\xi)] = E[\exp(\mathrm{i}\cdot\omega\cdot\xi_1)\cdot\exp(\mathrm{i}\cdot\omega\cdot\xi_2)].$$

$$E[\exp(\mathrm{i}\cdot\omega\cdot\xi)] = E[\exp(\mathrm{i}\cdot\omega\cdot\xi_1)]\cdot E[\exp(\mathrm{i}\cdot\omega\cdot\xi_2)],$$

  i.e., $\chi_\xi(\omega) = \chi_{\xi_1}(\omega)\cdot\chi_{\xi_2}(\omega)$.
- Hence, $\ln(\chi_\xi(\omega)) = \ln(\chi_{\xi_1}(\omega)) + \ln(\chi_{\xi_2}(\omega))$.
- So, if we define n-the logarithmic moment as $L_n(\xi) \stackrel{\text{def}}{=} \dfrac{1}{\mathrm{i}^n}\cdot\dfrac{d^n\chi_\xi(\omega)}{d\omega^n}|_{\omega=0}$, we conclude that

$$L_n(\xi) = L_n(\xi_1) + L_n(\xi_2).$$

- *Comment.* The factor $\dfrac{1}{\mathrm{i}^n}$ is added to make the moments real numbers.

8. Reconstructing x From y: Main Idea

- Since $y(t)=\sum_s a(s)\cdot x(t-s)$, we get

$$L_n(y)=\left(\sum_s (a(s))^n\right)\cdot L_n(x).$$

- *Idea:* so, we can reconstruct $L_n(x):=L_n(y)\Big/\left(\sum_s (a(s))^n\right)$.
- In the *ideal non-noise case*, once we know the exact distribution for y, we can reconstruct the desired distribution for x as follows:
  - first, we compute the logarithmic moments $L_n(y)$ of the signal y;
  - then, we compute the value $L_n(x)$;
  - finally, we use the Taylor series to reconstruct the logarithm of the characteristics function as

$$\ln(\chi_x(\omega))=L_1\cdot \mathrm{i}\cdot\omega+L_2\cdot \mathrm{i}^2\cdot\omega^2+L_3\cdot \mathrm{i}^3\cdot\omega^3+\ldots$$

So, we can determine the characteristic function $\chi_x(\omega)$ of the original distribution x.

9. From Characteristic Function to, e.g., Probability Density Function

- It is known that the characteristic function *uniquely* determines the distribution.
- For example, from its definition, we can describe its relation to the probability density function $\rho(x)$ as follows:

$$\chi(\omega) = E[\exp(\mathrm{i} \cdot \omega \cdot \xi)] = \int \exp(\mathrm{i} \cdot \omega \cdot x) \cdot \rho(x)\, dx.$$

- So, $\chi(x)$ is a *Fourier transform* of the probability density function.
- Hence, the original probability density function $\rho(x)$ can be determined as the *inverse Fourier transform* of the characteristics function

$$\rho(x) = \frac{1}{2\pi} \cdot \int \exp(-\mathrm{i} \cdot \omega \cdot x) \cdot \chi(\omega)\, d\omega.$$

10. Computations: Comments

- For each filter, computing $\sum_s (a(s))^n$ may be *computationally intensive*, but we must do it only once.
- The above description referred to the idealized *no-noise* case.
- In reality, the noise is always present: the whole purpose of the filter was to decrease this noise.
- Because of the noise, in practice, we can only reconstruct a few first logarithmic moments
$$L_1(x), L_2(x), \ldots, L_n(x).$$
- These moments do not determine the distribution uniquely:
- There exist several different distributions with the same values of the first moments.
- We must therefore select a distribution with given values of these logarithmic models.

11. From Logarithmic Moments L_n to Traditional Moments M_n

- The problem of reconstructing a distribution from the logarithmic moments is reasonably new.
- We will see that this problem is closely related to the well-studied problem:
  - reconstructing a distribution
  - from the standard moments

$$M_n \stackrel{\text{def}}{=} E[\xi^n] = \int x^n \cdot \rho(x)\,dx.$$

- We will show that:
  - knowing the first n logarithmic moments $L_1, \ldots, L_n$
  - is equivalent to knowing the first moments $M_1, \ldots, M_n$.

12. From L_n to M_n: cont-d

- By definition, $L_1 = \frac{1}{\mathrm{i}} \cdot \frac{\partial \ln(\chi)}{\partial \omega}_{|\omega=0}$. Using the chain rule for differentiation, we conclude that $\frac{\partial \ln(\chi)}{\partial \omega} = \frac{1}{\chi} \cdot \frac{\partial \chi}{\partial \omega}$. For $\omega = 0$, we have $\chi(0) = E[\exp(\mathrm{i} \cdot 0 \cdot \omega)] = 1$, and

$$\frac{\partial \chi}{\partial \omega}_{|\omega=0} = \left(\frac{\partial}{\partial \omega} E[\exp(\mathrm{i} \cdot \omega \cdot \xi)]\right)_{|\omega=0} =$$

$$E\left[\left(\frac{\partial}{\partial \omega} \exp(\mathrm{i} \cdot \omega \cdot \xi)\right)_{|\omega=0}\right] = E[\mathrm{i} \cdot \xi] = \mathrm{i} \cdot M_1.$$

  Therefore, $L_1 = \frac{1}{\mathrm{i}} \cdot \mathrm{i} \cdot M_1$, i.e., $L_1 = M_1$.

- Similarly, $L_2 = \frac{1}{\mathrm{i}^2} \cdot \frac{\partial^2 \ln(\chi)}{\partial \omega^2}_{|\omega=0}$, where

$$\frac{\partial^2 \ln(\chi)}{\partial \omega^2} = \frac{1}{\chi} \cdot \frac{\partial^2 \chi}{\partial \omega^2} - \frac{1}{\chi^2} \cdot \left(\frac{\partial \chi}{\partial \omega}\right)^2.$$

  For $\omega = 0$, we have $\frac{\partial^2 \chi}{\partial \omega^2}_{|\omega=0} = \mathrm{i}^2 \cdot M_2$, hence $L_2 = M_2 - M_1^2$, i.e., L_2 is the variance.

- Similarly, $L_3 = M_3 + 2M_1^3 - 2M_1 \cdot M_2$, etc.

13. Maximum Entropy Approach

- *Problem:* there are many possible probability distributions with the given values of the moments $M_1, \ldots, M_n$.

- *Possible answer:* $-\int \rho(x) \cdot \ln(\rho(x))\, dx \to \max\limits_{\rho}$ under the conditions $\int x^i \cdot \rho(x)\, dx = M_i$.

- *Solution:* Lagrange multipliers lead to

$$J = -\int \rho(x) \cdot \ln(\rho(x))\, dx + \sum_{i=0}^{n} \lambda_i \cdot \int x^i \cdot \rho(x)\, dx;$$

  differentiating w.r.t. ρ, we get

$$-\ln(\rho(x)) - 1 + \sum_{i=0}^{n} \lambda_i \cdot x^i = 0,$$

  i.e., $\rho(x) = C \cdot \exp(-\lambda_1 \cdot x - \ldots - \lambda_n \cdot x^n)$.

- *Algorithm:* C and λ_i can be determined from the fact that the overall probability should be 1, and the moments are M_i.

+ For $n = 2$, we get Gaussian distribution.

− For $n > 2$, a computationally intensive problem.

14. Alternatives to MaxEnt

- *Problem:* the Maximum Entropy approach is computationally intensive.
- *In the symmetric case* ($L_3 = 0$), use Weibull distributions

$$\rho(x) = \text{const} \cdot \exp(-k \cdot |x - a|^p).$$

+ These distributions indeed well describe measurement errors, e.g., in geophysics.

- *Algorithm* for finding p:

$$\varepsilon = \frac{\Gamma(1/p) \cdot \Gamma(5/p)}{\Gamma(3/p)^2}, \text{ where } \varepsilon \stackrel{\text{def}}{=} \frac{M_4}{M_2^2}.$$

- *Even faster approximate algorithm:*

$$p = \frac{1.46}{\ln(\varepsilon - 2/9 - 10.7/\varepsilon^7) - 0.289}.$$

- *In asymmetric case:*
  - $\rho(x) = \text{const}_- \cdot \exp(-k_- \cdot |x - a|^p)$ for $x \le a$ and
  - $\rho(x) = \text{const}_+ \cdot \exp(-k_+ \cdot |x - a|^p)$ for $x \ge a$.

15. Other Alternatives to the Maximum Entropy Approach

- *Fact:* previously, we required that for the first 2 moments, we get Gaussian distribution.
- *Idea:* for the first four moments, we can get even faster computations if we do not make this requirement.
- *Example:* Generalized Lambda distributions, in which the quantile function $Q(u)$ – inverse to the cumulative distribution function $F(t)$ – has the form

$$Q(u) = \lambda_1 + \frac{1}{\lambda_2} \cdot \left[\frac{u^{\lambda_3} - 1}{\lambda_3} - \frac{(1-u)^{\lambda_4} - 1}{\lambda_4} \right].$$

- To find λ_i, we solve a system of 2 non-linear equations with 2 unknowns.
- In these equations, non-linearity is also described by known special functions: namely, by beta functions.

16. The Use of Expert Knowledge

- Often, in addition to the four (or more) moments, we also have some expert knowledge about $\rho(x)$.
- This expert knowledge usually comes in terms of words from natural language.
- So, it is natural to use *fuzzy* techniques: $\mu(\rho)$.
- *Idea:* select ρ for which $\mu(\rho) \to \max$.
- *Comment:* this idea is used in geosciences (Bardossy, Demicco, Fodor, Klir, et al.).
- In the absence of additional expert information, this approach leads:
  - either to the Maximum Entropy formulas
  - or to *generalized entropy* $\int \rho(x)^\alpha \, dx$.

  In this case, $\rho(x) = (\lambda_0 + \lambda_1 \cdot x + \ldots + \lambda_n \cdot x^n)^{-\beta}$.

17. Acknowledgments

This work was supported in part:

- by NASA under cooperative agreement NCC5-209,
- by NSF grants EAR-0112968, EAR-0225670, and EIA-0321328,
- by NIH grant 3T34GM008048-20S1, and
- by ARL grant DATM-05-02-C-0046.