Speaker Locations from Inter-Speaker Range Measurements: Closed-Form Estimator and Performance Relative to the Cramèr-Rao Lower Bound

Jeff Walters, Scott Wilson, Jonathan Abel

{jhw, rswilson, abel}@ccrma.stanford.edu

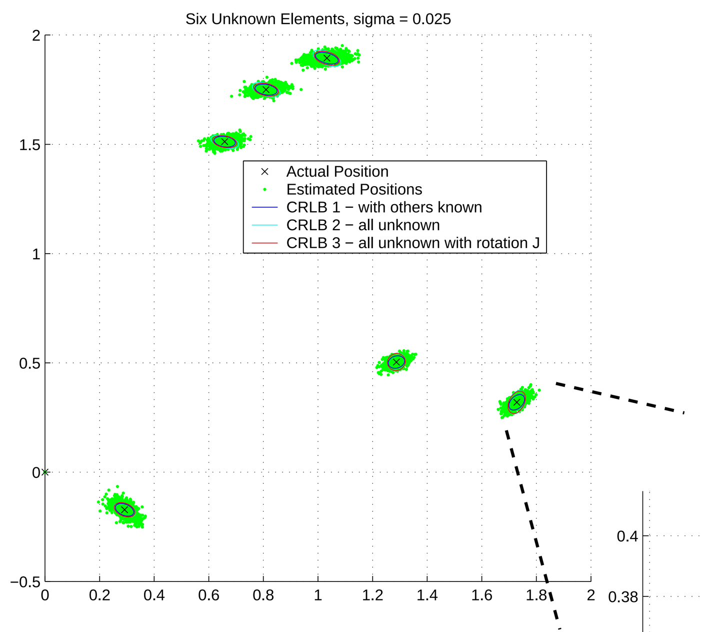
The problem of determining the positions of speakers in an array given noisy measurements of inter-speaker ranges is considered. A closed-form position estimator which minimizes a weighted equation error norm is presented. The information inequality is used to bound position estimate mean square error and to gauge the accuracy of the closed-form estimator, which is shown to be nearly efficient.

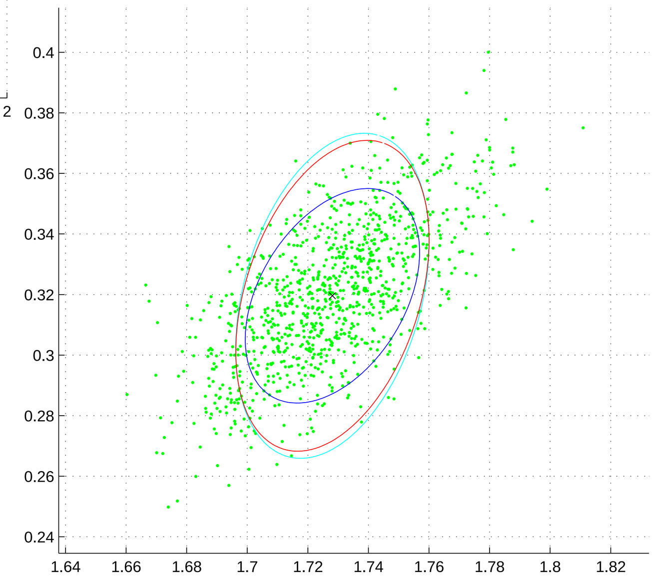
1. Introduction

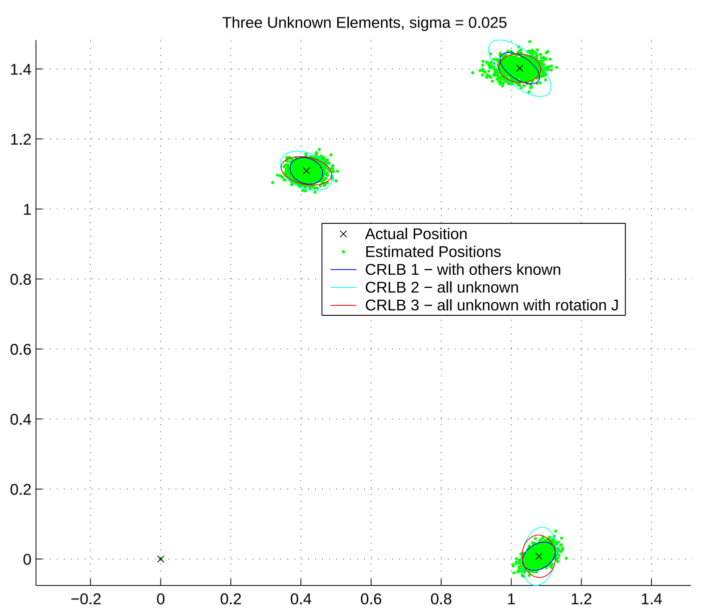
Many current techniques for multichannel rendering such as Ambisonics, VBAP, and wavefield synthesis are benefited by or are dependent upon accurate speaker location information. We study the problem of determining the positions of multiple speakers using noisy measurements of inter-speaker ranges.

Estimating the position of an object from range measurements to a set of fixed positions - the problem of intersecting spheres - arises in navigation systems. One estimation approach is to form an equation error relating the unknown position and range measurements, selected such that its norm is easily minimized and it produces hypothesized ranges close the measured values when its norm is small.

Simulation results: The speaker position estimator was applied to many realizations of the noisy inter-element range estimates. 95% error ellipses are plotted for the three CRLB formulations described below.







2. Position Estimation

Let $\mathbf{X}$ be the $N \times P$ matrix of speaker locations where N is the number elements and P is the dimensionality of the problem

$$\mathbf{X} \triangleq \begin{bmatrix} \mathbf{x}_0^\top \\ \vdots \\ \mathbf{x}_{N-1}^\top \end{bmatrix}.$$

The actual inter-element ranges r_{ij} are given by

$$r_{ij} = \| \mathbf{x}_i - \mathbf{x}_j \| . \tag{1}$$

We assume that the measured inter-element ranges ρ_{ij} are corrupted with additive, independent Gaussian noise with known variance

$$\rho_{ij} = r_{ij} + \epsilon_{ij} \sim \mathcal{N}(r_{ij}, \sigma^2), \tag{2}$$

as would be the case using any number of inter-speaker signal arrival time estimates. Under this assumption, the estimation problem becomes one of finding the parameters determining the mean of a Gaussian-distributed measurement.

In this case, the maximum likelihood estimate is known to be efficient (unbiased with minimum variance) for small measurement errors. The maximum likelihood estimate minimizes the weighted sum of square hypothesized measurement errors (differences between measured and hypothesized ranges), but is difficult to compute directly.

Stacking squared instances of (1), and taking $\mathbf{x}_0 = 0$, we have

$$2\mathbf{X}\mathbf{X}^\top = (\rho^2)\mathbf{1}^\top + \mathbf{1}(\rho^2)^\top - \mathbf{R} + \epsilon_{EE}, \tag{3}$$

where (ρ^2) is a column of square range measurements, $\mathbf{R}$ is a matrix of square range measurements, and ϵ_{EE} is an equation error.

We estimate the positions of the speaker elements $\hat{\mathbf{X}}$ to within an orthogonal transformation $\mathbf{Q}$ by constructing a matrix using the P largest singular values of the singular value decomposition of $\mathbf{X}\mathbf{X}^\top$. This matrix is known to minimize the sum of square equation error elements over all position sets X in P dimensions

$$\hat{\mathbf{X}} = \mathbf{U} \cdot \mathbf{D}^{\frac{1}{2}} \cdot \mathbf{Q}. \tag{4}$$

Noting that $\mathbf{U} \cdot \mathbf{D} \cdot \mathbf{V}^\top = \mathbf{X}\mathbf{X}^\top$ is the singular value decomposition of $\mathbf{X}\mathbf{X}^\top$.

3. Information Inequality

The information inequality states that the variance of any unbiased estimator will be greater than or equal to the inverse of the Fisher information, also known as the Cramèr-Rao Lower Bound (CRLB) of the estimator

$$var\{\hat{\mathbf{X}}\} \geqq \mathbf{J}_{\hat{\mathbf{X}}}^{-1}.$$

The bound is useful in gauging the performance of the equation error minimizer, and in developing insight into the information contained in the range measurements.

The Fisher information is given by

$$\mathbf{J}_{\hat{\mathbf{X}}} = \frac{\partial \mu^\top}{\partial \hat{\mathbf{X}}} \cdot \mathbf{\Sigma}_R^{-1} \cdot \frac{\partial \mu}{\partial \hat{\mathbf{X}}^\top}, \tag{5}$$

with μ defined as the collection of all inter-element range measurements

$$\mu^\top = [\cdots r_{ij} \cdots] .$$

Defining the unit vector pointing from element i to element j as β_{ij}

$$\beta_{ij} \triangleq \frac{\mathbf{x}_i - \mathbf{x}_j}{\|\mathbf{x}_i - \mathbf{x}_j\|}, \tag{6}$$

the sensitivity of the mean to changes in the ith element position is

$$\frac{\partial \mu^\top}{\partial \mathbf{x}_i} = [0 \cdots 0 \, \beta_{ij} \, 0 \cdots 0] .$$

When all but one of the speaker positions are known, the Fisher Information is the outer product of direction vectors pointing from the unknown speaker to the other speakers, weighted by the respective inverse measurement variances:

$$\mathbf{J} = \mathbf{B}^\top \mathbf{\Sigma}^{-1} \mathbf{B} = \sum_{i=0}^{N-1} \frac{1}{\sigma_i^2} \beta_i \beta_i^\top. \tag{7}$$

This interpretation is consistent with the notion that accurate estimates are possible along directions having a large number of elements.

When there are two unknown speaker positions ($\mathbf{x}_0$ and $\mathbf{x}_1$) we write the Fisher information $\tilde{\mathbf{J}}_0$ as the sum of the information contained in the ranges measured to the known speakers $\mathbf{J}_0$ and the information in the range measured between the unknown elements:

$$\tilde{\mathbf{J}}_0 = \mathbf{J}_0 + \frac{\beta_{01}\beta_{01}^\top}{\sigma_{01}^2 + \beta_{01}^\top \mathbf{J}_1^{-1} \beta_{01}}. \tag{8}$$

Note that the additional information provided by the measurement relative to the unknown position takes on a form similar to that of a term in (7). It is as if the second unknown position were known, but its associated range measurement variance were increased.

The CRLB for speaker element $\mathbf{x}_0$ is

$$var\{\hat{\mathbf{x}}_0\} \geqq \mathbf{J}_0^{-1} - \frac{\mathbf{J}_0^{-1}\beta_{01}\beta_{01}^\top \mathbf{J}_0^{-1}}{\sigma_{01}^2 + \beta_{01}^\top (\mathbf{J}_0^{-1} + \mathbf{J}_1^{-1}) \beta_{01}}. \tag{9}$$

The variance decrease is large when β_{01} aligns with a major axis of the previous error ellipse.

4. Observations

- As the number of elements increases, the bound for the case of all element positions unknown approaches that of a single position unknown.
- As the elliptical bounds become more flattened, we see a greater spread between the bounds for the case of all element positions unknown versus that of a single position unknown.
- Arrays with elements grouped in a line result in higher variance perpendicular to that line.
- The Fisher information for the rotational ambiguity of the array estimate noticeably effects the bound variance only for arrays of a small number of elements.
- The estimator seems to be performing at approximately 1.5 times the theoretical MVU estimator.