STATISTICAL CHALLENGES IN MODERN ASTRONOMY IV
ASP Conference Series, Vol. 371, 2007
G. Jogesh Babu and Eric D. Feigelson

Nonparametric Estimation of Dark Matter Distributions

Xiao Wang

Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD 21250, USA

Michael Woodroofe and Jayanta Pal

Department of Statistics, University of Michigan, Ann Arbor, MI 48109, USA

Matthew Walker and Mario Mateo

Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA

Abstract. There have been recent advances in the nonparametric estimation of the distribution of dark matter in nearby dwarf spheroidal galaxies. This work uses Jeans' Equations to relate the distribution of (total) matter to velocity dispersions. The estimation of velocity dispersions is facilitated by observing certain shape restrictions, stronger that convexity in the isotropic case, that follow from Jeans' Equations. This work is reviewed, modified slightly, and extended to include some anisotropy.

Keywords: Dwarf spheroidal galaxies; Jeans' Equations; Shape restricted estimation; Wicksell's problem.

1. Introduction

The dwarf spheroidal galaxies near to the Milky Way are among the least luminous objects in the night sky. While they have stellar population sizes similar to those of globular clusters, approximately $10^6 - 10^7$ stars, they are spread out over a much larger area, $2 - 6$ kpc. In many cases it is unclear how these galaxies could have avoided tidal disruption by the Milky Way without the presence of considerable unseen matter. In this vein Richstone and Tremaine (1986) have proposed models which imply that the dwarf spheroidals are dominated by unseen matter, suggesting that the dwarf spheroidals may be excellent candidates for studying dark matter. Early studies of dark matter distributions relied on a few ($\sim$ 30) stars, and researchers were forced to adopt highly structured parametric models, like King's model. See Mateo (1993). Recent advances in instrumentation have changed this dramatically. The MIKE Fiber Optic System now makes it possible to obtain velocity measurements on up to 256 stars in approximately the same time earlier researchers required for one. Samples of size 1000-1500 are now obtainable and, so, nonparametric statistical analyses are possible.

Wang et al. (2005) and Walker et al. (2006) have developed some nonparametric techniques for estimating dark matter distributions. Assuming spherical symmetry and isotropy, they use Jeans' Equations, (5) and (6) below, to relate the distribution of mass to velocity dispersions and then show how to estimate the velocity dispersions. From Jeans' Equations, it develops that a function that must be estimated as an intermediate step is convex, and this fact is essential to the estimation process. This work is reviewed here, modified slightly, and extended to allow some anisotropy.

Notation is established in Section 2. Estimation of the spatial distribution of stars is considered in Section 3. This problem has similar features to the estimation of velocity dispersions, but is simpler. Unsurprisingly, the isotropic case is simplest. This case is considered in Section 4, and the method is explained there. Extensions to some anisotropic cases are outlined in Section 5. These indicate that the estimate of total mass can be vastly greater in the presence of moderate anisotropy, but not substantially smaller. Some practical complications are considered in Section 6.

2. Preliminaries

Consider a single dwarf spheroidal galaxy. Let $\boldsymbol{x} = (x_1, x_2, x_3)$ denote position within the galaxy so normalized that $\mathbf{0} = (0,0,0)$ is the center, and let $r = \|\boldsymbol{x}\| := \sqrt{x_1^2 + x_2^2 + x_3^2}$ be distance from the center. Suppose throughout that the density of (total) mass is of the form $\rho(\boldsymbol{x}) = \rho_0(r)$. If ρ were known then the movements of the stars in the galaxy could be determined from the potential function

$$\Phi(\boldsymbol{x}) = \int_{\mathbb{R}^3} \frac{G\rho(\boldsymbol{y})d\boldsymbol{y}}{\|\boldsymbol{x} - \boldsymbol{y}\|} = \Phi_0(r), \text{ say},$$

which also depends on $\boldsymbol{x}$ only through r. The problem faced here is the inverse problem: From the movement of the stars, determine ρ. Let

$$M(r) = 4\pi \int_0^r t^2 \rho_0(t) dt \tag{1}$$

denote the mass within r units from the center. The statistical goal is to estimate the functions $M(r)$ and $\rho_0(r)$.

To estimate $M(r)$ we suppose that a sample of n stars has been observed. For the present, we suppose that this is a random sample; selection effects will be considered in Section 6. Let $\boldsymbol{X} = (X_1, X_2, X_3)$ denote the position of a randomly selected star and $\boldsymbol{V} = (V_1, V_2, V_3)$ its velocity; let $f(\boldsymbol{x}, \boldsymbol{v})$ denote the joint probability density of $\boldsymbol{X}$ and $\boldsymbol{V}$, so that

$$P[\boldsymbol{x} \le \boldsymbol{X} \le \boldsymbol{x} + d\boldsymbol{x}, \boldsymbol{v} \le \boldsymbol{V} \le \boldsymbol{v} + d\boldsymbol{v}] = f(\boldsymbol{x}, \boldsymbol{v}) d\boldsymbol{x} d\boldsymbol{v};$$

and suppose throughout that $\boldsymbol{V}$ has been centered, so that

$$E(\boldsymbol{V}) = \int_{\mathbb{R}^6} \boldsymbol{v} f(\boldsymbol{x}, \boldsymbol{v}) d\boldsymbol{x} d\boldsymbol{v} = 0. \tag{2}$$

Here we use the statistical notation $E(\boldsymbol{V})$ for expectation; astronomers might write this as $\langle \boldsymbol{V} \rangle$.

Represent $\boldsymbol{x}$ in spherical coordinates as $x_1 = r\sin(\theta)\cos(\phi)$, $x_2 = r\sin(\theta)\sin(\phi)$, and $x_3 = r\cos(\theta)$, where $r = \sqrt{x_1^2 + x_2^2 + x_3^2}$, $-\pi < \phi = \arctan(x_2/x_1) \le \pi$ and $0 \le \theta = \arccos(x_3/r) < \pi$; and let

$$\begin{aligned} v_r &= \sin(\theta)\cos(\phi)v_1 + \sin(\theta)\sin(\phi)v_2 + \cos(\theta)v_3, \\ v_\phi &= -\sin(\phi)v_1 + \cos(\phi)v_2, \\ v_\theta &= \cos(\theta)\cos(\phi)v_1 + \cos(\theta)\sin(\phi)v_2 - \sin(\theta)v_3, \end{aligned}$$

and $\boldsymbol{w} = (v_r, v_\phi, v_\theta)$. For a fixed $\boldsymbol{x}$, this transformation is orthogonal, and its inverse is $v_1 = \sin(\theta)\cos(\phi)v_r - \sin(\phi)v_\phi + \cos(\theta)\cos(\phi)v_\theta$, $v_2 = \sin(\theta)\sin(\phi)v_r + \cos(\phi)v_\phi + \cos(\theta)\sin(\phi)v_\theta$, and $v_3 = \cos(\theta)v_r - \sin(\theta)v_\theta$.

Suppose now that the joint density of position $\boldsymbol{X}$ and velocity $\boldsymbol{V}$ has the form

$$f(\boldsymbol{x}, \boldsymbol{v}) = f_0(r^2, \boldsymbol{w}).$$

Then the marginal density of $\boldsymbol{X}$, $\int_{\mathbb{R}^3} f(\boldsymbol{x}, \boldsymbol{v}) d\boldsymbol{v}$, has the form

$$\int_{\mathbb{R}^3} f_0(r^2, \boldsymbol{w}) d\boldsymbol{w} = f_0(r^2), \text{ say.}$$

In Jeans' Equations, an important role is played by the radial velocity dispersion for which the statistical notation is $\sqrt{E(V_r^2|\boldsymbol{X} = \boldsymbol{x})} = \sigma_r(r)$, say. Here

$$E(V_r^2|\boldsymbol{X} = \boldsymbol{x}) = \frac{\nu_r(r)}{f_0(r^2)} \tag{3}$$

where

$$\nu_r(r) = \int_{R^3} v_r^2 f(\boldsymbol{x}, \boldsymbol{v}) d\boldsymbol{v} = \int_{\mathbb{R}^3} v_r^2 f_0(r^2, \boldsymbol{w}) d\boldsymbol{w}.$$

We will be able to estimate this radial velocity dispersion from the sample. The corresponding quantities, $\nu_\phi(r) = \int_{R^3} v_\phi^2 f(\boldsymbol{x}, \boldsymbol{v}) d\boldsymbol{v}$ and $\nu_\theta(r) = \int_{R^3} v_\theta^2 f(\boldsymbol{x}, \boldsymbol{v}) d\boldsymbol{v}$ are similarly related to conditional expectation but are less useful.

The form of Jeans' Equations used below allows for some anisotropy. Specifically, suppose that $f_0(r^2, \boldsymbol{w})$ depends on $\boldsymbol{w}$ only through $v_\phi^2 + v_\theta^2$ and v_r and that

$$\nu_\theta = (1 - beta)\nu_r, \tag{4}$$

where $-\infty < \beta < 1$ is a constant, an anisotropy parameter. As $\beta \to -\infty$, the motion of the stars approaches a pure rotation, with very small radial velocity; $\beta = 0$ corresponds to isotropy; and as β approaches 1 the stars move back and forth through the center with only a slight rotation. Then

$$\nu_r(r) + 2\beta\frac{\nu_r}{r} + f_0(r^2)\Phi_0'(r) = 0, \tag{5}$$

and

$$M(r) = -\frac{r^2}{G f_0(r^2)}\left[\nu_r'(r) + 2\beta\frac{\nu_r(r)}{r}\right] \tag{6}$$

by Jeans' Equations, Equations (4.54) and (4.55) of Binney and Tremaine (1987) in a different notation. [The $\nu(r)$ in Binney and Tremaine (1987) corresponds to our $f_0(r^2)$]. Thus, $M(r)$ can be simply expressed in terms of f_0 and ν_r, and the problem is to estimate these quantities. Of course, we do not observe $\boldsymbol{X}$ and $\boldsymbol{V}$ completely, but only the velocity in the line of sight (from earth) and the projection of $\boldsymbol{X}$ on the plane orthogonal to the line of sight. With a proper choice of coordinate axes, these become X_1, X_2, and V_3; and the problems become those of estimating the functions f_0 and ν_r, defined in three dimensions, from lower dimensional data. The techniques for solving these two problems are similar, but the first is simpler. It is considered next.

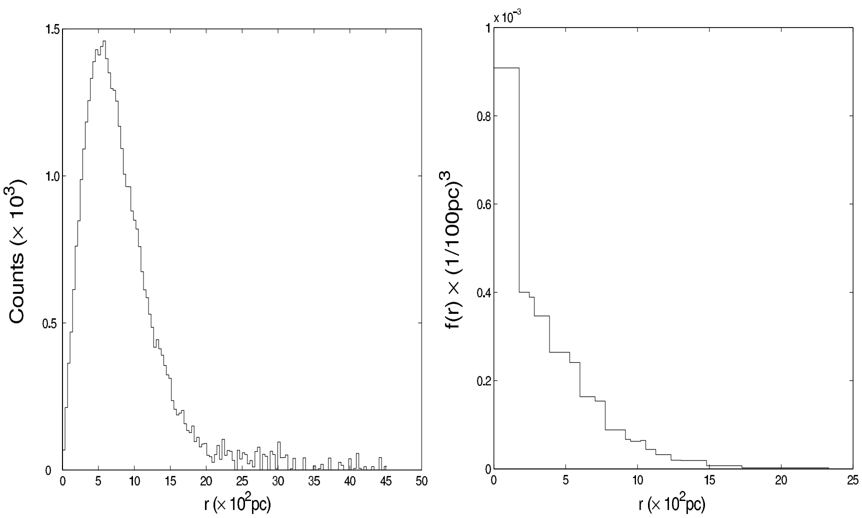

Figure 1. The left panel is the projected counts data for Fornax in Irwin and Hatzidimitriou (1995). The right is the estimated three-dimensional distribution f_0 of stars for Fornax assuming spherical symmetry.

3. Estimating f_0

Fortunately, Irwin and Hatzidimitriou (1995) have provided very large data sets from which f_0 can be estimated. For each of nine dwarf spheroidals, the data consists of counts of the number of stars whose projections fall in circular annuli; that is, letting $0 = r_0 < r_1 < \cdots < r_m < \infty$ be the radii of the concentric circles, the data count the number of stars for which $r_{k-1} < \sqrt{X_1^2 + X_2^2} \leq r_k$ for $k = 1, \cdots, m$. Table 1 indicates the nature of the data. The problem is to turn these two dimensional data into an estimate of the density $f_0(r^2)$ of $\boldsymbol{X}$ in three dimensions. This problem is known as Wicksell's (1925) Problem. Our solution follows Groeneboom and Jongbloed (1995).

Table 1. Star Counts From Fornax

r	d	N		r	d	N
1.05	29.33	97	$\cdots$	110.16	1.29	0
2.10	30.57	304	$\cdots$	111.21	1.42	59
3.15	31.31	519	$\cdots$		$\cdots$	
		$\cdots$	$\cdots$	$\cdots$		
27.28	9.14	1277		139.11	1.34	0

Notes: The r's are the radii of the concentric circle in arcmin; the d's are the densities of background and Fornax stars in counts/arcmin^2; the background density is 1.34; N is the number of Fornax stars in the annuli.

Let g denote the joint density of X_1 and X_2, and $y = x_1^2 + x_2^2$, the squared radius of the projection. Then

$$g(x_1, x_2) = \int_{-\infty}^{\infty} f_0(r^2)dx_3 = \int_y^{\infty} \frac{f_0(z)dz}{\sqrt{z-y}} = g_0(y), \tag{7}$$

say. For example, if f_0 is a step function, say $f_0(z) = \phi_k$ for $r_{k-1}^2 < z \leq r_k^2,\ k = 1, \cdots, m$, then

$$g_0(y) = 2\phi_k\sqrt{r_k^2 - y} + 2\sum_{j=k+1}^{m} \phi_j \left[\sqrt{r_j^2 - y} - \sqrt{r_{j-1}^2 - y}\right]$$

for $r_{k-1}^2 < y \leq r_k^2$.

For purposes of estimation, we will suppose that f_0 is a step function and estimate the ϕ_k's. If f_0 is a step function and a star is selected at random, then the probability that its projection falls outside the inner $i-1$ annuli is

$$p_i = \pi \int_{r_{i-1}^2}^{\infty} g_0(y)dy = \sum_{j=i}^{m} a_{ij}\phi_j, \tag{8}$$

where

$$a_{ij} = \frac{4\pi}{3}\left[\sqrt{(r_j^2 - r_{i-1}^2)^3} - \sqrt{(r_{j-1}^2 - r_{i-1}^2)^3}\right]$$

for $j \geq i$, and $a_{ij} = 0$ for $j < i$. Letting N denote the total number of stars in sample and N_i the number in the ith annulus, $p_i^{\#} = (N_i + \cdots + N_m)/N$ provides an unbiased estimator of p_i for each i. Then, solving equation (8) with p_i replaced by $p_i^{\#}$ yields unbiased estimators of ϕ_i,

$$\phi_i^{\#} = \frac{1}{a_{ii}}\left[p_i^{\#} - \sum_{j=i+1}^{m} a_{ij}\phi_j^{\#}\right], \tag{9}$$

where the sum is to be interpreted as 0 when $i = m$. There are two possible problems with (9): $\phi_i^{\#}$ may be unstable for small i, and (9) can produce negative estimates. The cure for both problems is to collapse some of the bins. If one of the $\phi_i^{\#}$'s is negative, simply combine that bin with one of the two adjacent ones, and continue the process until all of the $\phi_i^{\#}$'s are non-negative. Similarly, if the estimate of $\phi_1^{\#}$ appears to be unreasonable, combine the first few bins. This introduces some limited subjectivity into the process.

Another option is to make some qualitative assumptions about f_0. If we suppose that f_0 is a decreasing function, then we could seek the values of ϕ_i that minimize

$$\int_0^{\infty} r^2[f_0(r^2) - f_0^{\#}(r^2)]^2 dr = \frac{1}{3}\sum_{i=1}^{m}(r_i^3 - r_{i-1}^3)(\phi_i - \phi_i^{\#})^2$$

among all non-increasing sequences $\phi_1 \geq \cdots \geq \phi_m$. There is an analytical expression for the solution:

$$\hat{\phi}_j = \min_{i \leq j} \max_{k \geq i} \frac{\sum_{l=j}^{k}(r_l^3 - r_{l-1}^3)\phi_l^{\#}}{r_k^3 - r_{j-1}^3}.$$

Using $\hat{\phi}_i$ in place of $\phi_i^{\#}$ produces non-negative estimates, but can still lead to instability for small i, and some collapsing of bins may still be necessary.

Figure 1 shows the projected counts data for Fornax in Irwin and Hatzidimitriou (1995) and the estimated three-dimensional distribution f_0 of stars for Fornax.

4. The Isotropic Case

The situation is simplest in the isotropic case, where $f(\boldsymbol{x}, \boldsymbol{v})$ depends only on r and $v = \sqrt{v_1^2 + v_2^2 + v_3^2}$ and $\beta = 0$. Recall that (only) the line of sight velocities are observed, and let

$$\nu_3(r) = \int_{\mathbb{R}^3} v_3^2 f(\boldsymbol{x}, \boldsymbol{v}) d\boldsymbol{v}.$$

Then $\nu_r(r) = \nu_3(r)$ in the isotropic case, and the missing V_1 and V_2 cause no problem. This is the most important simplification that comes with isotropy. The missing X_3 does cause a problem, however. Let $y = x_1^2 + x_2^2$ and

$$\psi(y) = \int_{-\infty}^{\infty} \int_{\mathbb{R}^3} v_3^2 f(\boldsymbol{x}, \boldsymbol{v}) d\boldsymbol{v} dx_3 = \int_{-\infty}^{\infty} \int_{\mathbb{R}^3} v_3^2 f_0(y + x_3^2, \boldsymbol{w}) d\boldsymbol{w} dx_3.$$

Then $E[V_3^2 | X_1 = x_1, X_2 = x_2] = \psi(y)/g_0(y)$ and

$$\psi(y) = \int_y^{\infty} \frac{\nu_3(\sqrt{z}) dz}{\sqrt{z - y}}, \tag{10}$$

as in (7). Let

$$\Psi(t) = \int_t^{\infty} \frac{\psi(y) dy}{\sqrt{y - t}}. \tag{11}$$

Then

$$\Psi(t) = \pi \int_t^{\infty} \nu_3(\sqrt{z}) dz \tag{12}$$

by reversing the orders of integration in (12),

$$\Psi'(t) = -\pi \nu_3(\sqrt{t}) \qquad \text{and} \qquad \Psi''(r^2) = \frac{G\pi f_0(r^2)}{2r^3} M(r) \tag{13}$$

from (6) with $\beta = 0$.

As in (9) there is a simple unbiased estimator of Ψ. Denote the sample by $(X_{i,1}, X_{i,2}, V_{i,3}),\ i = 1, \cdots, n$. Let $Y_i = X_{i,1}^2 + X_{i,2}^2$ and

$$\Psi^{\#}(t) = \frac{1}{n} \sum_{i: Y_i > t} \frac{V_{i,3}^2}{\sqrt{Y_i - t}}. \tag{14}$$

Then $\Psi^{\#}(t)$ is an unbiased estimator of $\Psi(t)$ for each t, but $\Psi^{\#}$ is highly irregular when viewed as a function of t with infinite discontinuities at each of the Y_i. This may be contrasted with Ψ, which must be convex since $\Psi''(t) > 0$ from (13) (and more). The

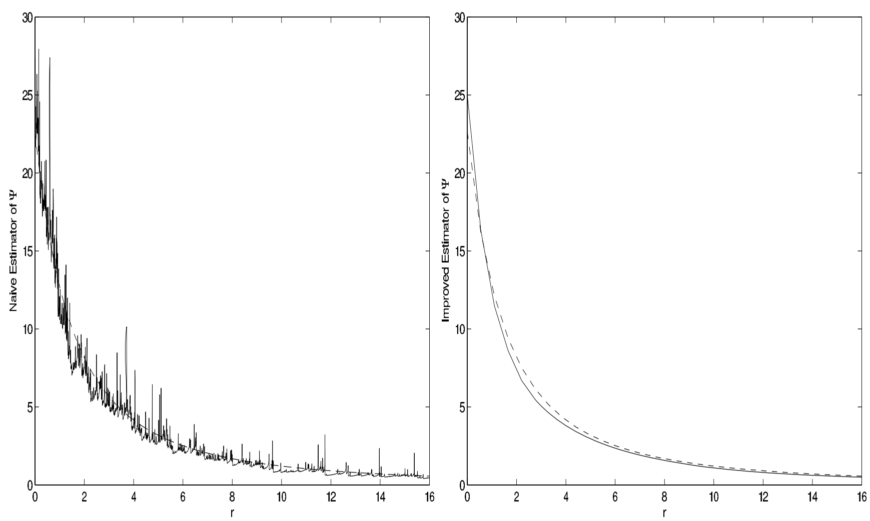

Figure 2. The function $\Psi^{\#}$ (left) and $\hat{\Psi}$ (right) for a simulated sample. The solid line is the estimate and the dashed line the true function.

irregularity can be cured by requiring the estimator to be convex, as explained below. The results are displayed in Figure 2.

We will model M as a slightly modified quadratic spline, and this will involve some binning of the data. Let $0 = r_0 < r_1 \cdots < r_m < \infty$ be constants (for example, from Irwin and Hatzidimitriou (1995)), and suppose that

$$M(r) = \theta_1 a(r) + \sum_{i=2}^{m} \frac{\theta_i}{2}(r - r_{i-1})_+^2 \tag{15}$$

where $(x)_+$ is the larger of x and zero, $(x)_+^2$ is the square of $(x)_+$, and

$$a(r) = \begin{cases} \frac{r^3}{3r_1} & r \le r_1 \\ \frac{1}{2}(r^2 - \frac{r_1^2}{3}) & r > r_1 \end{cases}$$

is an adjustment included to force $M(r)$ to be a multiple of r^3 on the interval $[0, r_1]$. Then $M'(r)$

is piecewise linear on the interval $[r_1, r_m]$ and quadratic on the interval $[0, r_1]$. It follows easily that a necessary and sufficient condition for M to be a bounded increasing function is that

$$\sum_{i=1}^{m} \theta_i (r_j - r_{i-1})_+ \ge 0 \tag{16}$$

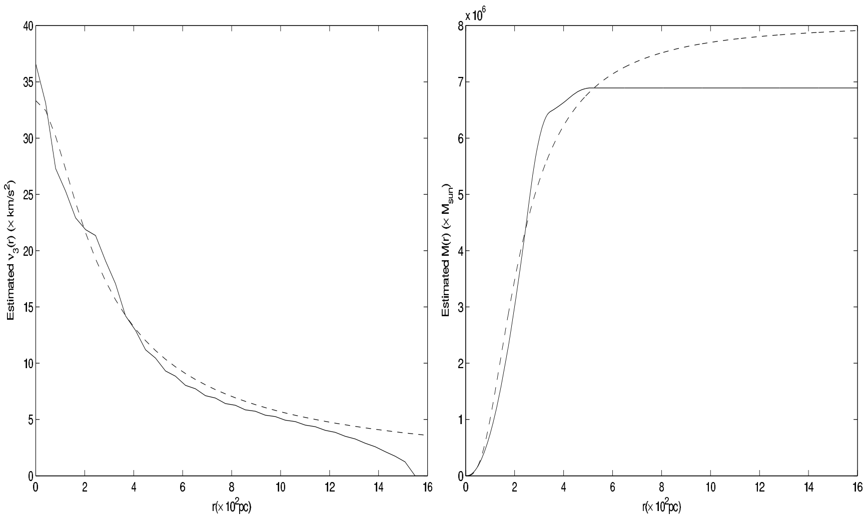

Figure 3. $\widehat{\nu_3}$ (left) and $\hat{M}$ (right) for a simulated sample. The solid line is the estimator, and the dashed line the true.

for $j = 1, \cdots, m$ with equality when $j = m$. Similarly, a necessary and sufficient condition for ρ to be decreasing is that

$$\sum_{i=1}^{k} \theta_i [2r_{i-1} - r_j] \leq 0 \tag{17}$$

for $j = k - 1$ and $j = k$ and $k = 1, \cdots, m$. Importantly, these are linear constraints of the θ_i.

If M is as in (15), then

$$\Psi(t) = \sum_{i=1}^{m} \theta_i \Gamma_i(t), \tag{18}$$

where

$$\Gamma_i(t) = \int_t^\infty (s-t) \frac{G\pi f_0(s)}{4\sqrt{s^3}} (\sqrt{s} - r_{i-1})_+^2 ds,$$

for $i = 2, \cdots, m$ and

$$\Gamma_1(t) = \int_t^\infty (s-t) \frac{G\pi f(s)}{2\sqrt{s^3}} \left[a(\sqrt{s})\right] ds$$

To impose the shape restrictions, it seems natural to minimize the integral of $[\Psi(t) - \Psi^{\#}(t)]^2$ over functions Ψ of the form (18), subject to (16) or both (16) and (17). Unfortunately, $\Psi^{\#}$ is so irregular that the integral of $(\Psi^{\#})^2$ is infinite. However, it would be equivalent to minimize

$$\kappa = \int_0^\infty \Psi(t)^2 dt - 2 \int_0^\infty \Psi(t)\Psi^{\#}(t) dt,$$

which is finite. The latter expression may be written

$$\kappa = \sum_{i=1}^{m} \sum_{j=1}^{m} q_{ij}\theta_i\theta_j - 2\sum_{i=1}^{m} z_i\theta_i, \tag{19}$$

where

$$z_i = \int_0^{r_m^2} \Psi^{\#}(t)\Gamma_i(t)dt$$

and

$$q_{ij} = \int_0^{r_m^2} \Gamma_i(t)\Gamma_j(t)dt.$$

The z_i and q_{ij} can be computed, by numerical integration if necessary. So, the problem of minimizing (19), subject to (16) or both (16) and (17), is a quadratic programming problem. Let $\hat{\theta}_1, \cdots, \hat{\theta}_m$ denote the solution to the quadratic programming problem. Then

$$\hat{\Psi}(t) = \sum_{i=1}^{m} \hat{\theta}_i\Gamma_i(t) \tag{20}$$

provides estimator of Ψ and that does satisfy the shape restrictions; and estimators of the radial velocity dispersion and mass profile may be obtained by differentiating (20), or by substituting $\hat{\theta}_i$ for θ_i in (15). The estimators are illustrated in Figure 3 for a simulated sample. The estimators appear quite reasonable except (perhaps) for large and small values of r. The reason for this problem is quite simple: there is not much data near the endpoints. See Figure 1.

The reader may well object that it would be more natural to model M as a cubic spline, instead of a modified quadratic one. The mathematical problem is that with cubic splines, the constraints on the θ_i are no longer linear, and the quadratic programming problem becomes a more difficult conic programming problem. This question is studied in some detail in Li and Wang (2006).

5. Anisotropy

If $\beta \neq 0$ in (4), then the situation is more complicated, but can still be analyzed. The major complication is that ν_r is no longer equal to ν_3. In the anisotropic case $\nu_3(r,\theta) = \int_{I\!R^3} v_3^2 f(\boldsymbol{x},\boldsymbol{v})d\boldsymbol{v}$ depends on both r and θ and

$$\begin{aligned}\nu_3(r,\theta) &= \int_{I\!R^3} \left[\cos(\theta)v_r - \sin(\theta)v_\theta\right]^2 f(r^2,\boldsymbol{w})d\boldsymbol{w} \\ &= \cos^2(\theta)\nu_r(r) + \sin^2(\theta)\nu_\theta(r).\end{aligned} \tag{21}$$

So, using (4),

$$\nu_3(r,\theta) = [1 - \beta\sin^2(\theta)]\nu_r(r). \tag{22}$$

Define ψ and Ψ as above. Then the relation between Ψ and ν_r is

$$\Psi(t) = \pi\int_t^{\infty}[1 - \beta\frac{t+z}{2z}]\nu_r(\sqrt{z})dz,$$

as in (12). So,

$$\Psi'(t) = -(1-\beta)\pi\nu_r(\sqrt{t}) - \frac{1}{2}\beta\pi\int_t^\infty \frac{\nu_r(\sqrt{z})}{z}dz,$$

and

$$\Psi''(r^2) = -\frac{1}{2r}(1-\beta)\pi\nu_r'(r) + \frac{1}{2}\beta\pi\frac{\nu_r(r)}{r^2} \tag{23}$$

Thus, ν_r can be recovered from Ψ'' by solving a simple differential equation, and then M can be computed from (6). Omitting some details, the relevant equation is

$$M(r) = \frac{2r^3}{G\pi(1-\beta)f_0(r^2)}\left[\Psi''(r^2) - (\alpha+\beta)r^{2\alpha-2}\int_{r^2}^\infty \frac{\Psi''(z)dz}{z^\alpha}\right],$$

where $\alpha = \beta/[2(1-\beta)]$.

The estimation of M and ν_r now proceeds as above. First $\Psi^\#$ in (14) is still an unbiased estimator; M is again modeled as a modified quadratic spline as in (15), and a relation of the form (18) holds, with a more complicated definition of Γ_i. The estimation of the θ_i is again reduced to a quadratic programming, and the estimation of Ψ and M proceeds as above. The results are displayed in Figure 4 for a data set comprising 179 Fornax Stars. Three values of β are included, representing isotropy and two cases of moderate anisotropy. Observe that the estimate of total mass can be vastly larger in the presence of moderate anisotropy, but not substantially smaller. The estimate of total mass is at least 3×10^8 solar masses, substantially larger than earlier estimates [Mateo (1993)]. The sample size 179 here is a bit too small, but help is on the way. Data have now been collected on ~ 1000 Fornax stars, though not yet reduced.

6. Complications

There are two complications that must be overcome before the procedure can be applied: velocities are observed with error and there may be selection effects.

For a sample, $(X_{i,1}, X_{i,2}, V_{i,3})$ of projected positions and line of sight velocities, suppose that $V_{i,3}$ is observed with error, say $U_i = V_{i,3} + \epsilon_i,\ i = 1, \cdots, n$, where $\epsilon_1, \cdots, \epsilon_n$ are random errors with known standard deviations $\sigma_1, \cdots, \sigma_n$. Then replacing $V_{i,3}^2$ by $U_i^2 - \sigma_i^2$ in (14) still produces an unbiased estimator of Ψ and the procedure described in the last section can proceed with only that minimal change.

Selection effects are more complicated. Suppose that the probability of observing a star with projected position (x_1, x_2) can depend (only) on x_1 and x_2, say $w_0(x_1, x_2)$. Then the joint probability density of X_1, X_2, and $\boldsymbol{V}$ for stars in the sample becomes

$$g^*(x_1, x_2, \boldsymbol{v}) = w(y)g_0(y, \boldsymbol{v}), \tag{24}$$

where $y = x_1^2 + x_2^2$,

$$w(y) = \frac{1}{c}\int_{-\pi}^{\pi} w_0[\sqrt{y}\sin(\phi), \sqrt{y}\cos(\phi)]d\phi,$$

and c is a normalizing constant. If an astronomer can specify the function w, then

$$\Psi^\#(t) = \frac{1}{n}\sum_{i:Y_i>t} \frac{U_i^2 - \sigma_i^2}{w(Y_i)\sqrt{Y_i - t}} \tag{25}$$

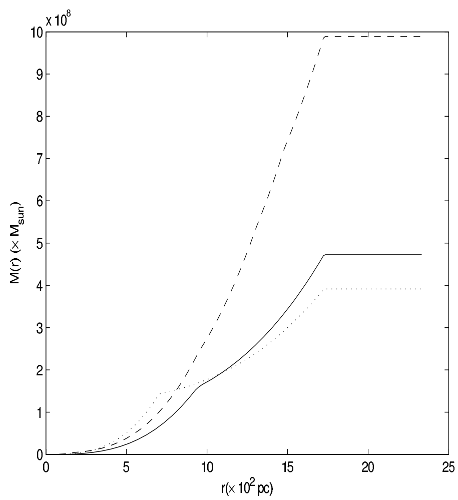

Figure 4. The function $\hat{M}$ for a sample of 179 stars from a Fornax. The solid line assumes that $\beta = 0$; the dashed and dotted lines that $\beta = -2$ and $\beta = .2$. The data are reported in detail by Walker et al. (2006).

provides an unbiased estimator of Ψ, and estimation ν_r and M can proceed as described in the previous section.

If the astronomer cannot specify w, then w can be estimated from the sample. Integrating over $\boldsymbol{v}$ in (24) shows that the probability density of X_1 and X_2 for stars in the sample is $g^*(x_1, x_2) = w(y)g_0(y) = g_0^*(y)$, say. So $w(y) = g_0^*(y)/g_0(y)$. If f_0 is known [from Irwin and Hatzidimitrou (1995), for example], then $g(x_1, x_2) = g_0(y)$, where

$$g_0(y) = \int_y^\infty \frac{f_0(z)dz}{\sqrt{z-y}}.$$

Further, g_0^* can be estimated from the stars in the sample. For example, using a kernel estimator,

$$\hat{g}_0^*(y) = \frac{1}{nh}\sum_{i=1}^n K\left(\frac{y - Y_i}{h}\right),$$

where K is a kernel and $h > 0$ is a bandwidth. Then

$$\hat{w}(y) = \frac{\hat{g}_0^*(y)}{g_0(y)}$$

provides an estimator of w which can be substituted into (25). Figures 2 and 4 use (25).

Acknowledgments. The research of all authors was supported by the National Science Foundation.

References

Binney, J. and Tremaine, S. (1987). *Galactic Dynamics*. Princeton Series in Astrophysics.

Groeneboom, P. and Jongbloed, K. (1995). Isotonic estimation and rates of convergence in Wicksell's Problem. *Annals of Statistics*, **23**, 1518-1542.

Irwin, M. and D. Hatzidimitriou (1995). Structure parameters for the galactic dwarf spheroidals. *MNRAS*, **277**, 1354-1378.

Li, F. and Wang, X. (2006). Shape restricted smoothing spline regression. Submitted.

Mateo, M. (1993). Dwarf galaxies of the local group. *Annu. Rev. Astron. Astrophy.*, **36**, 435-506.

Richstone, D.O. and Tremaine, S. (1986). Measuring mass-to-light ratios of spherical stellar systems by core fitting. *Astronomical Journal*, **92**, 72-74.

Walker, M., Mateo, M., Olzewski, E., Bernstein, R., Wang, X. and Woodroofe, M. (2006). Internal kinematics of the Fornax Dwarf Spheroidal Galaxy. To appear in *The Astronomical Journal*.

Wang, X., Woodroofe, M., Walker, M., Mateo, M. and Olshzewski, E. (2005). Estimating dark matter distributions. *The Astrophysical Journal*, **626**, 145-158.

Wicksell, S.D. (1925). The corpuscular problem. *Biometrika*, **17**, 84-99.