

Modelling stochastic changes in curve shape, with an application to cancer diagnostics

Asger Hobolth and Eva B. Vedel Jensen

University of Aarhus

Abstract

Often, the statistical analysis of the shape of a random planar curve is based on a model for a polygonal approximation to the curve. In the present paper, we describe instead the curve as a continuous stochastic deformation of a template curve. The advantage of this continuous approach is that the parameters in the model do not relate to a particular polygonal approximation. A somewhat similar approach has been used in Kent *et al.* (1996), who describe the limiting behaviour of a model with a first-order Markov property as the landmarks on the curve become closely spaced, see also Grenander (1993). The model studied in the present paper is an extension of this model. Our model possesses a second-order Markov property. Its geometrical characteristics are studied in some detail and an explicit expression for the covariance function is derived. The model is applied to the boundaries of profiles of cell nuclei from a benign tumour and a malignant tumour. It turns out that the model with the second-order Markov property is the most appropriate, and that it is indeed possible to distinguish between the two samples.

1 Introduction

In the grading of malignancy of cancer tissue, many morphological parameters may be used. At low magnification the architecture of the cellular tissue

Postal address: Laboratory for Computational Stochastics, Department of Mathematical Sciences, University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark. Email addresses: asho@imf.au.dk and ev@imf.au.dk.

may be considered while on a medium scale of magnification the variation in cell nuclear size is an important feature. At high magnification the shape, size and colour of each single nuclei are studied. Stereological techniques may be used to determine the size of nuclei, cf. e.g. Sørensen (1991) and Jensen (1998). The remaining parameters are, however, usually subjectively estimated by the pathologists without using any quantitative methods. The pathologists' opinions may differ, and particularly in the cases where the malignancy is in an intermediate stage this leads to different grading. In order to make precise diagnoses there is therefore a need for supplementary methods, which may objectively quantify important features at each scale of magnification. This paper presents a method of describing stochastic changes of the shape of solid objects in the plane with no obvious landmarks. The method is applied to nuclear profiles, obtained by sectioning normal tissue and cancer tissue from the human skin.

The basic idea is to model an observed planar curve $\mathcal{F} = \{F(t) : 0 \leq t \leq T\}$, which is the boundary of a solid object, as a stochastic deformation of a non-random closed template curve $\mathcal{C} = \{c(t) : 0 \leq t \leq T\}$. One of the stochastic geometry models considered takes the form

$$F(t) = c(t) + X(t)\omega(t), \quad 0 \leq t \leq T,$$

where $\omega(t)$ is the inner unit normal vector to \mathcal{C} at $c(t)$, cf. Figure 1. Note that $X(t)$ is the signed distance between $F(t)$ and $c(t)$, and may therefore be regarded as a residual. The challenge is to model the residual process $\{X(t) : 0 \leq t \leq T\}$.

We will assume that the residual process is distributed as

$$\{TX_1(t/T) : 0 \leq t \leq T\},$$

where the distribution of the 'normalized' process $\{X_1(t) : 0 \leq t \leq 1\}$ belongs to a parametrized class $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ of distributions of cyclic and stationary stochastic processes with zero mean. Since the curves considered are closed, the residual process should be cyclic. In the application, the residual process is assumed to be Gaussian with zero mean and with a second-order Markov property. Recall that a Gaussian process $\{X(t)\}$ with zero mean is stationary if and only if the covariance between $X(t)$ and $X(s)$ depends on (t, s) through $s \Leftrightarrow t$ only. Stationarity is a natural requirement when the object has no obvious landmarks and hence no reference point.

Note that under translations, rotations and rescaling in the plane, the normalized process $\{X_1(t) : 0 \leq t \leq 1\}$ remains unchanged. Therefore, this

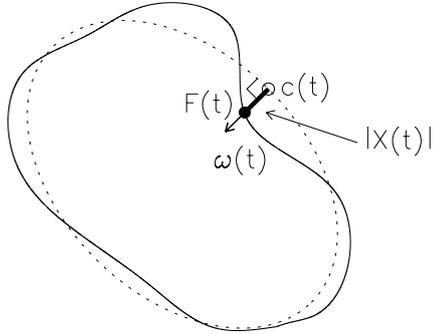


Figure 1: The observed curve is a realisation of a stochastic deformation of a template.

process describes the stochastic changes in curve shape of the observed curve \mathcal{F} relative to the template curve \mathcal{C} , and the changes can be quantified by estimating the parameter of the distribution P_θ of the normalized process. One of the major advantages of this continuous type model is its independence from the need to specify the number of landmarks.

The idea of describing objects such as potatoes, cells, hands or leaves as deformations of a template has been advocated by Ulf Grenander. His work on pattern theory has been collected in Grenander (1993). In the above mentioned examples the template is a closed polygon, representing the outline of a typical object. Grenander and Manbeck (1993) use a discretized ellipse with fixed eccentricity as template in an application concerning defect detection in potatoes. In order to determine whether an object has the shape of a potato, the angles of the edges of the discretized object are compared with the angles of the edges of the template. In our model a similar comparison naturally occurs when we consider the derivative of the residual process.

In Rue and Syversveen (1998), a procedure for identifying cells in a digital image is developed. Their prior model is an example of a stochastic geometry model. The template curve \mathcal{C} is a circle with a radius r , and the residual process

$$\{X(t) : 0 \leq t \leq 2\pi r = T\}$$

is a cyclic and stationary Gaussian process with zero mean, variance $r^2\sigma^2$

and correlation function

$$\rho(h) = \begin{cases} e^{-\alpha h/T} \cos(4\pi h/T), & 0 \leq h \leq T/2, \\ \rho(T \Leftrightarrow h), & T/2 \leq h \leq T, \end{cases}$$

where $\alpha > 0$.

Kent *et al.* (1996) consider multivariate normal models for edges and vertices of a closed polygonal outline in the plane. The inverse covariance matrix is a circular matrix with a first-order Markov property. They describe the limiting behaviour when the vertices become closely spaced. We extend this approach to the case where the inverse covariance matrix is a circular matrix with a second-order Markov property. This turns out to be a better choice for our purpose. The general second-order model, described in the Appendix of the present paper, contains as a special case a second-order model suggested by Grenander (1993).

Kass *et al.* (1988) provide through the theory of snakes a way of performing boundary detection in an image. A snake is a curve in an image which minimizes a certain energy functional. In our set-up \mathcal{F} is the snake while \mathcal{C} can be regarded as a template snake. A very first choice of the energy functional could take the form

$$E_{snake}^* = \int_0^T (\alpha X(t)^2 + \beta X'(t)^2 + E_{image}(t)) dt, \quad \alpha, \beta > 0. \quad (1)$$

The first term represents an external constraint energy, which forces the snake to have a shape similar to the template snake. The next term represents an internal energy and makes the snake smooth. The last term connects the snake to the image. A simple image energy functional could be

$$E_{image}(t) = \delta I(F(t)), \quad 0 \leq t \leq T,$$

where $I(x)$ is the intensity of the image at the position x . Depending on the sign of δ the snake is attracted to either black or white pixels. This approach is very similar to a Bayesian algorithm for object detection with the first-order Markov model described in Section 4.1 below as the prior model.

In Section 2 the data is described, and an ellipse is fitted to each of the objects. In Section 3 the stochastic geometry models are presented, and some geometrical characteristics of the residual process are explored under these models. In Section 4 we consider first- and second-order Markov models for the residual process, and they are fitted to the data in Section 5. Finally in Section 6 we discuss further properties of the models considered in Section 4 and consider topics for future work.

2 The data

The data set consists of 27 nuclear profiles from a malignant tumour and 27 nuclear profiles from a benign tumour of the human skin. The silhouettes were drawn by hand directly from the microscope screen. The data have previously been analysed with respect to size and variability of size in Jensen and Sørensen (1991) and Sørensen (1991).

By visual inspection the nuclear profiles were smoothed and rescaled such that an area of approximately 75,000 pixels was obtained for each profile, cf. Figure 2. As may be seen from Figure 2, nuclear profiles from the malignant tumour appear to be less smooth than those from the benign tumour.

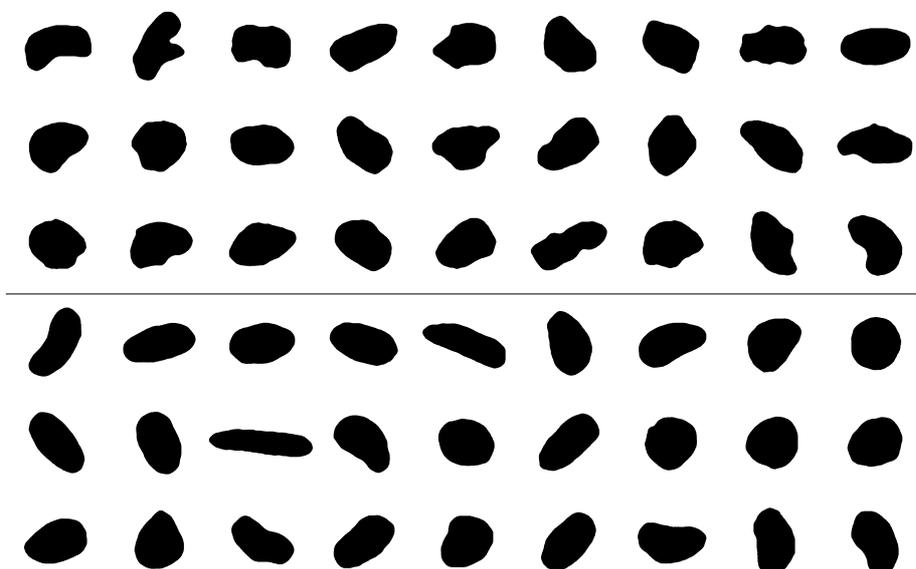


Figure 2: The nuclear profiles after scaling and smoothing. The upper panel is from the malignant tumour while the lower is from the benign tumour.

Prior to any analysis it is necessary to identify the boundary of each nuclear profile and to generate an ordered list of pixels on the boundary and their positions. We will now describe the algorithm we used for this job. A boundary pixel is a profile pixel having as one of its 8 neighbours a pixel from outside the profile. The algorithm starts at one of the boundary pixels and searches for another in the directions shown in Figure 3. Having found a

boundary pixel in direction i , say, the algorithm proceeds as follows. If a new boundary pixel is found in direction i , then direction i is maintained as the search direction. Otherwise direction $(i+1 \bmod 16)$ is applied, and if it leads to a boundary pixel, it is the new search direction. Otherwise direction $(i-1 \bmod 16)$ is applied etc., cf. Figure 3. This procedure led to around 350-400 ordered boundary pixels for each profile.

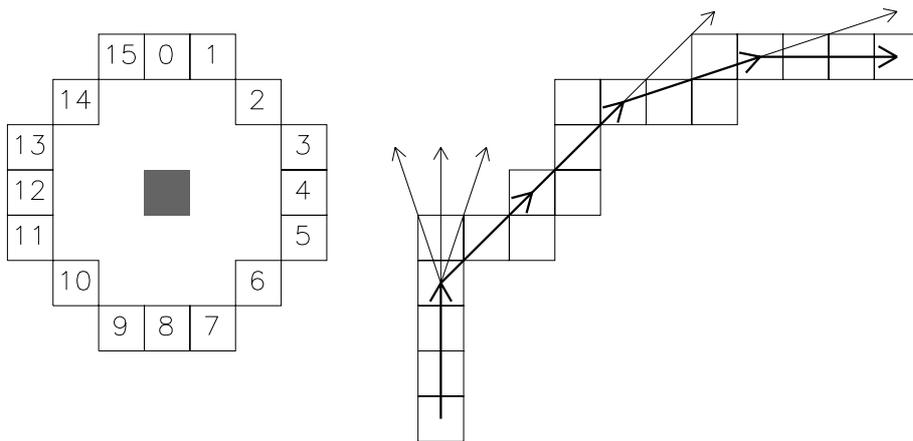


Figure 3: The 16 search directions and an example of the algorithm.

As a first analysis we fitted an ellipse to each of the profiles, using $n = 50$ (approximately) equidistant points $(x_1, \dots, x_n) \in \mathbb{R}^{2n}$ on the boundary of the profile. Let $\mathcal{C} = \mathcal{C}(a, b, \theta, x_0, y_0)$ be an ellipse with semi-axes a, b , orientation θ and center (x_0, y_0) , and denote by $c_i(a, b, \theta, x_0, y_0)$ the point on \mathcal{C} closest to x_i , $i = 1, \dots, n$, cf. Figure 4. The fitted ellipse was then determined as that having parameter values (a, b, θ, x_0, y_0) that minimizes

$$\sum_{i=1}^n |x_i \ominus c_i(a, b, \theta, x_0, y_0)|^2,$$

where $|\cdot|$ is Euclidean norm. The minimum can be found by standard numerical recipes. We also experimented with values of n larger than 50 and observed only minor changes in the estimated parameter values.

In Figure 5 we have plotted the ratio between the minor axis and the major axis of the fitted ellipses of the profiles. It is clear, that even though the ratios of the benign tumour profiles appear to have a wider range, there is no significant difference between the two samples. Therefore, a study of

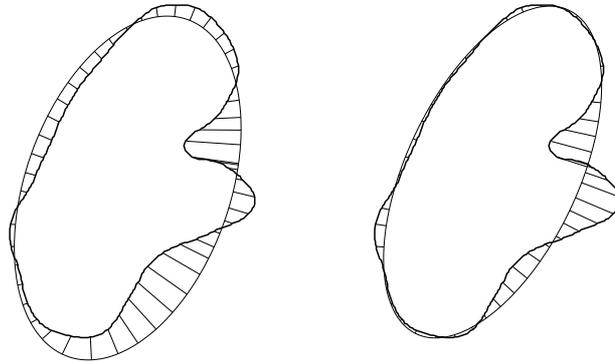


Figure 4: The fitted ellipse minimizes the sum of squares of the indicated distances. Left: An initial fit. Right: The fitted ellipse.

the residual process is needed in order to describe the difference in shape of the two samples.

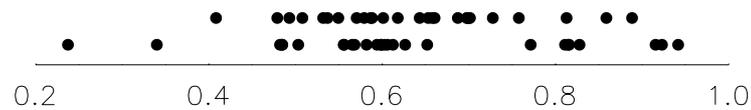


Figure 5: The ratio between the minor axis and the major axis of the fitted ellipses of the profiles. The upper points are from the profiles from the malignant tumour while the lower points are from the profiles from the benign tumour.

3 Stochastic geometry models for random curves

When modelling featureless objects using templates some of the frequently posed questions concern the choice, matching and number of landmarks. We propose a model where the observed curve \mathcal{F} is matched to the template curve \mathcal{C} by a perpendicular projection. A continuous type description makes it possible to study how the parameters of the finite-dimensional distributions

depend on the number of landmarks. There are several ways of choosing the landmarks. In Model I below we parametrize \mathcal{C} by arc length, which in the application corresponds to choosing the landmarks equidistantly on \mathcal{C} . In Model II we take the reverse approach and parametrize \mathcal{F} by arc length. Thinking of \mathcal{F} as a deformation of \mathcal{C} , Model I is perhaps the most appropriate. One disadvantage of Model I is, however, that curve segments from \mathcal{F} which are close to being perpendicular to \mathcal{C} are undersampled. This is obviously not the case for Model II.

3.1 Model I

The observed curve $\mathcal{F} = \{F(t) \in \mathbb{R}^2 : 0 \leq t \leq T\}$ is assumed to be a realization of a stochastic process

$$F(t) = c(t) + X(t)\omega(t), \quad 0 \leq t \leq T. \quad (2)$$

Here, $\mathcal{C} = \{c(t) \in \mathbb{R}^2 : 0 \leq t \leq T\}$ is a non-random closed ($c(0) = c(T)$) smooth curve in the plane parametrized by arc length. Furthermore, $\omega(t)$ is the inner unit normal vector to \mathcal{C} at $c(t)$, $0 \leq t \leq T$, and

$$\{X(t) \in \mathbb{R} : 0 \leq t \leq T\}$$

is a real-valued cyclic and stationary stochastic process with zero mean. The process $\{X(t)\}$ models the deviations between the observed curve \mathcal{F} and the expected curve \mathcal{C} . We will call $\{X(t)\}$ the residual process.

Note that the construction (2) puts some restrictions on how 'wild' the random curve \mathcal{F} may look. Thus, each point $c(t) \in \mathcal{C}$ generates exactly one point on the random curve, positioned on the line $c(t) + \text{span}\{\omega(t)\}$. In particular, if \mathcal{C} is a circle with center e and radius r and $X(t) \leq r$, then \mathcal{F} will be the boundary of a random set which is star-shaped relative to e . In Figure 6, an example of a curve, which cannot be generated by (2), is shown.

The model is closed under translations and rotations in \mathbb{R}^2 . Under such transformations, the curve \mathcal{C} will be translated and rotated correspondingly, while $X(t)$ is unaffected. The model is also closed under scale transformations in \mathbb{R}^2

$$(x, y) \rightarrow \alpha(x, y) = (\alpha x, \alpha y), \quad \alpha > 0.$$

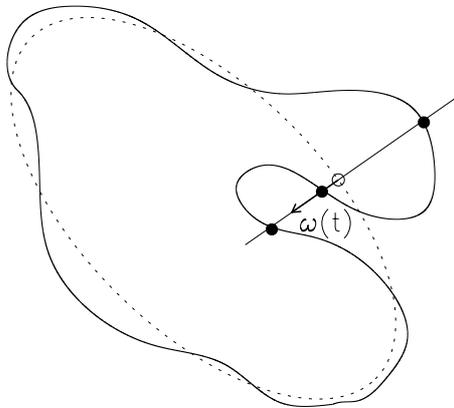


Figure 6: An example of a curve, which cannot be generated by (2).

Thus, parametrizing the rescaled curve $\alpha\mathcal{C}$ by arc length and letting $c_\alpha(t) = \alpha c(t/\alpha)$, we have

$$\alpha\mathcal{C} = \{c_\alpha(t) : 0 \leq t \leq \alpha T\}.$$

Furthermore, if we likewise define F_α and X_α , but let $\omega_\alpha(t) = \omega(t/\alpha)$, the equation for the scale-transformed process becomes

$$F_\alpha(t) = c_\alpha(t) + X_\alpha(t) \cdot \omega_\alpha(t), \quad 0 \leq t \leq T_\alpha,$$

where $T_\alpha = \alpha T$.

The features of the model (2) which are invariant under changes in location, orientation and scale are thus the shape of \mathcal{C} and the distribution of the process

$$X_1(t) = X(Tt)/T, \quad 0 \leq t \leq 1.$$

The latter process will be called the normalized residual process.

In order to analyze the model (2) we need to find, for selected values of $t \in [0, T]$, the point $F(t)$ on the line $c(t) + \text{span}\{\omega(t)\}$. If the fluctuations of $F(t)$ around $c(t)$ are not too large, then $F(t)$ is expected to be the point on $c(t) + \text{span}\{\omega(t)\}$ which is nearest to $c(t)$. In Figure 7, an example of too large fluctuations is shown.

The process $\{X(t)\}$ and its derivatives (if they exist) contain interesting geometric information about the difference between the random curve $F(t)$ and its expectation $c(t)$.

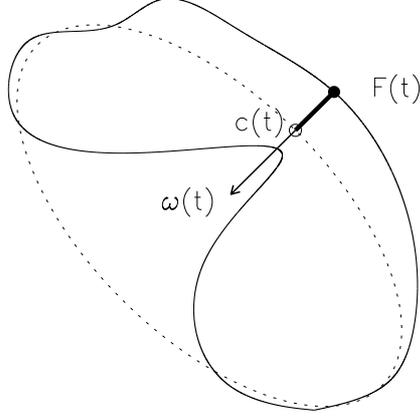


Figure 7: The fluctuations of $F(t)$ around $c(t)$ are large.

Proposition 1 *Suppose that $X(t)$ is differentiable. Let \mathcal{C} be orientated anti-clockwise. Furthermore, let $\Phi_F(t)$ and $\Phi_C(t)$ be the angles that $F'(t)$ and $c'(t)$ make with a fixed axis, respectively. Then,*

$$X'(t) = \tan \Psi(t)(1 \Leftrightarrow \kappa_C(t)X(t)), \quad (3)$$

where $\Psi(t) = \Phi_F(t) \Leftrightarrow \Phi_C(t)$ and $\kappa_C(t)$ is the curvature of \mathcal{C} at $c(t)$.

Proof. Using that

$$\begin{aligned} F(\tilde{t}) &= F(t) + F'(t)(\tilde{t} \Leftrightarrow t) + o_F(\tilde{t} \Leftrightarrow t) \\ c(\tilde{t}) &= c(t) + c'(t)(\tilde{t} \Leftrightarrow t) + o_C(\tilde{t} \Leftrightarrow t) \end{aligned}$$

we find that

$$\begin{aligned} 2X(t)X'(t) &= (X(t)^2)' \\ &= \lim_{\tilde{t} \rightarrow t} \frac{X(\tilde{t})^2 \Leftrightarrow X(t)^2}{\tilde{t} \Leftrightarrow t} \\ &= 2 < F(t) \Leftrightarrow c(t), F'(t) \Leftrightarrow c'(t) >, \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ is Euclidean inner product. Therefore,

$$\begin{aligned}
X'(t) &= \left\langle \frac{F(t) \Leftrightarrow c(t)}{X(t)}, F'(t) \Leftrightarrow c'(t) \right\rangle \\
&\stackrel{(\star)}{=} \left\langle \frac{F(t) \Leftrightarrow c(t)}{X(t)}, F'(t) \right\rangle \\
&= \left\langle \frac{F(t) \Leftrightarrow c(t)}{X(t)}, \frac{F'(t)}{|F'(t)|} \right\rangle |F'(t)|,
\end{aligned} \tag{4}$$

where at (\star) we have used that $c'(t) \perp F(t) \Leftrightarrow c(t)$.

Since \mathcal{C} is parametrized by arc length, $|c'(t)| = 1$ and

$$c'(t) = (\cos \Phi_{\mathcal{C}}(t), \sin \Phi_{\mathcal{C}}(t)).$$

It follows that

$$c''(t) = \Phi'_{\mathcal{C}}(t)(\Leftrightarrow \sin \Phi_{\mathcal{C}}(t), \cos \Phi_{\mathcal{C}}(t)) = \kappa_{\mathcal{C}}(t)\omega(t).$$

Using that

$$\frac{F'(t)}{|F'(t)|} = (\cos \Phi_F(t), \sin \Phi_F(t))$$

we therefore find

$$\left\langle \frac{F(t) \Leftrightarrow c(t)}{X(t)}, \frac{F'(t)}{|F'(t)|} \right\rangle = \sin(\Phi_F(t) \Leftrightarrow \Phi_{\mathcal{C}}(t)).$$

On the other hand,

$$\left\langle c'(t), \frac{F'(t)}{|F'(t)|} \right\rangle = \cos(\Phi_F(t) \Leftrightarrow \Phi_{\mathcal{C}}(t)) \tag{5}$$

and rewriting the left-hand side of (5)

$$\begin{aligned}
\left\langle c'(t), \frac{F'(t)}{|F'(t)|} \right\rangle &= \frac{1}{|F'(t)|} (\langle c'(t), c'(t) \rangle + \langle c'(t), F'(t) \Leftrightarrow c'(t) \rangle) \\
&= \frac{1}{|F'(t)|} (1 \Leftrightarrow \langle c''(t), F(t) \Leftrightarrow c(t) \rangle) \\
&= \frac{1}{|F'(t)|} (1 \Leftrightarrow \kappa_{\mathcal{C}}(t)X(t)),
\end{aligned}$$

(3) follows immediately. \square

Note that if the variance of $X(t)$ is small then $\Phi_F(t) \Leftrightarrow \Phi_C(t)$ is also expected to be small and

$$X'(t) \approx \Phi_F(t) \Leftrightarrow \Phi_C(t)$$

can be approximated by the process of angular differences. These differences have been considered in a discrete set-up by Grenander and Manbeck (1993). If $X(t)$ is twice differentiable, then under the same assumption

$$X''(t) \approx \Phi'_F(t) \Leftrightarrow \Phi'_C(t) = |F'(t)|\kappa_F(t) \Leftrightarrow \kappa_C(t) \approx \kappa_F(t) \Leftrightarrow \kappa_C(t).$$

3.2 Model II

As an alternative, we may take a reverse approach and parametrize \mathcal{F} by arc length. The idea is then to construct \mathcal{F} from a residual process $\{X(t) : 0 \leq t \leq T\}$ such that $X(t)$ is the signed distance from $F(t)$ to \mathcal{C} and the parameter t represents arc length on \mathcal{F} . It is evidently also necessary in order to start the construction to specify a point $c \in \mathcal{C}$ such that $F(0)$ has signed distance $X(0)$ to \mathcal{C} .

Notice that it is not always possible to construct \mathcal{F} in this way. As a simple example, suppose that \mathcal{C} is a circle of radius r . Then, to avoid pathological cases we must assume $X(t) \leq r$. Furthermore, for small $\epsilon > 0$, we also need

$$|X(t + \epsilon) \Leftrightarrow X(t)| \leq \epsilon,$$

cf. Figure 8. In particular, if $X(t)$ is differentiable, then $|X'(t)| \leq 1$.

Note that as under Model I, the 'shape features' of Model II are the shape of \mathcal{C} and the distribution of the normalized process $\{X(Tt)/T : 0 \leq t \leq 1\}$.

Similarly to the case where \mathcal{C} is parametrized by arc length, the process $X'(t)$ can be approximated by the process of angular differences.

Proposition 2 *Suppose that $X(t)$ is differentiable. Let \mathcal{C} be orientated anti-clockwise. Furthermore, let $\Phi_F(t)$ and $\Phi_C(t)$ be the angles that $F'(t)$ and $c'(t)$ make with a fixed axis, respectively. Then,*

$$X'(t) = \sin \Psi(t),$$

where $\Psi(t) = \Phi_F(t) \Leftrightarrow \Phi_C(t)$.

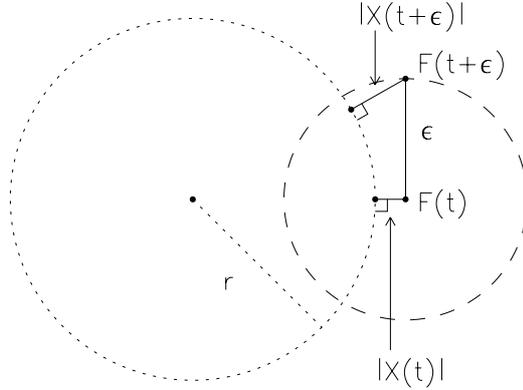


Figure 8: Illustration of the condition on $X(t)$.

Proof. Since \mathcal{F} is parametrized by arc length, $|F'(t)| = 1$ and

$$F'(t) = (\cos \Phi_F(t), \sin \Phi_F(t)).$$

Using that

$$\frac{c'(t)}{|c'(t)|} = (\cos \Phi_c(t), \sin \Phi_c(t)) \perp \omega(t)$$

we see that

$$\omega(t) = (\Leftrightarrow \sin \Phi_c(t), \cos \Phi_c(t)).$$

The proposition now follows from (4). □

4 Statistical inference

We will assume that the normalized residual process $\{X_1(t) : 0 \leq t \leq 1\}$ has a distribution belonging to a parametrized class of cyclic and stationary Gaussian processes with zero mean. In what follows we will omit the index 1 which should not cause any confusion.

There are as many choices of classes of Gaussian processes as there are parametrized classes of covariance functions. We will here concentrate on classes, having the property that the finite-dimensional (multivariate normal) distributions have a simple parametric form.

4.1 First-order Markov model

One such class which has been suggested, among others, by Grenander (1993, p. 476) and Kent *et al.* (1996) has the property that the finite-dimensional distributions have a first-order Markov property, approximately.

This class is characterized by the fact that $X(t)$ has zero mean, variance τ^2 and correlation function

$$\rho(h) = \rho(X(t), X(t+h)) = \frac{e^{(h-1/2)\phi} + e^{-(h-1/2)\phi}}{e^{\phi/2} + e^{-\phi/2}}, \quad 0 \leq h \leq 1, \quad \phi > 0. \quad (6)$$

Note that the correlation $\rho(h)$ is always positive and it is a decreasing function on the interval $[0, 1/2]$. Note also that $\rho(h) = \rho(1 \leftrightarrow h)$ which is a general property of cyclic and stationary processes on $[0, 1]$. The minimal correlation is

$$\rho(1/2) = \frac{2}{e^{\phi/2} + e^{-\phi/2}}$$

and can range from 0 to 1, cf. Figure 9.

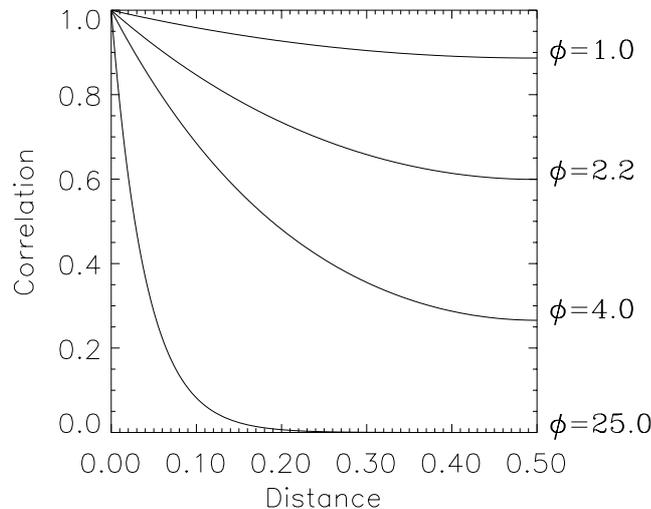


Figure 9: Correlation functions for the first-order Markov model for different values of ϕ .

In the proposition below, it is shown that the finite-dimensional distributions of $\{X(t)\}$ have a first-order Markov property, approximately. One of

the reasons why this result is important is that then the finite-dimensional distributions have a simple form, approximately, and from the corresponding likelihood functions it is easy to estimate the parameters of the model.

Recall that a multivariate normally distributed random vector

$$(X_0, \dots, X_{n-1}) \sim N_n(0, \Sigma) \quad (7)$$

is said to be Markov with respect to a graph with vertices $\{0, 1, \dots, n \Leftrightarrow 1\}$ if for $i, j \in \{0, 1, \dots, n \Leftrightarrow 1\}$

$$i \not\sim j \Rightarrow X_i \perp X_j | \{X_k : k \neq i, j\}. \quad (8)$$

Here, $i \not\sim j$ indicates that i and j are not neighbours in the graph and the right-hand side of (8) indicates that X_i and X_j are conditionally independent given the remaining coordinates, cf. Lauritzen (1996). Furthermore, it can be shown that the conditional independence in (8) is valid if and only if $(\Sigma^{-1})_{ij} = 0$, cf. Lauritzen (1996, p. 129). The random vector (7) is said to have a first-order Markov property if it is Markov with respect to the cyclic graph, which is given by

$$i \sim j \Leftrightarrow i = j \Leftrightarrow 1, j + 1 \pmod n. \quad (9)$$

We are now ready to formulate the proposition. For a $n \times n$ circular matrix $\{b_{ij}\}_{i,j=0}^{n-1}$ we use the notation $\text{circl}(a_0, a_1)$ if

$$b_{ij} = \begin{cases} a_0 & i = j \\ a_1 & i = j \Leftrightarrow 1, j + 1 \pmod n \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 3 *Let $\{X(t)\}$ be a cyclic and stationary Gaussian process with zero mean, variance τ^2 and correlation function (6). Furthermore, let $\{X^{(n)}(t)\}$ be the cyclic Gaussian process defined by*

$$X^{(n)} = (X^{(n)}(t_0), \dots, X^{(n)}(t_{n-1})) \sim N_n(0, \Sigma_n),$$

with $t_i = i/n, i = 0, \dots, n \Leftrightarrow 1$, with linear interpolation between t_i and t_{i+1} , and with $X^{(n)}(0) = X^{(n)}(1)$. Here, Σ_n is the regular $n \times n$ matrix with inverse

$$\Sigma_n^{-1} = \text{circl}(\alpha/n + 2\beta n, \Leftrightarrow \beta n), \quad \alpha, \beta > 0. \quad (10)$$

Then, $X^{(n)}$ is first-order Markov and $\{X^{(n)}(t)\}$ converges weakly to $\{X(t)\}$. The 1-1 correspondence between (τ^2, ϕ) and (α, β) is

$$\phi^2 = \alpha/\beta \quad \tau^2 = \frac{e^{\phi/2} + e^{-\phi/2}}{2\phi\beta(e^{\phi/2} \Leftrightarrow e^{-\phi/2})}.$$

A proof of Proposition 3 may be found in Grenander (1993, p. 476-480). It follows immediately from the form of Σ_n^{-1} that $X^{(n)}$ is first-order Markov.

It can be shown, using Lauritzen (1996, (5.11)), that

$$\begin{aligned} & \rho(X^{(n)}(t_i), X^{(n)}(t_j) | X^{(n)} \setminus \{X^{(n)}(t_i), X^{(n)}(t_j)\}) \\ &= \begin{cases} 1 & i = j \\ \frac{1}{(\phi/n)^2 + 2} & i = j \Leftrightarrow 1, j + 1 \pmod n \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

The parameters τ^2 and ϕ contain, under the model given in Proposition 3, the information about the deviation of the random curve shape from \mathcal{C} , cf. Section 3. The parameter τ^2 is a measure of the overall difference while ϕ regulates the smoothness of the random curve. The smaller the value of ϕ , the smoother the curve, cf. Figure 10.

Estimation of the parameters (τ^2, ϕ) or equivalently (α, β) can be done by an approximate likelihood analysis. Let

$$X_n = (X(t_0), \dots, X(t_{n-1})).$$

Then, because of Proposition 3, we can approximate the distribution of X_n with that of $X^{(n)}$ and use the simpler approximate likelihood

$$L_n(\alpha, \beta) = (2\pi)^{-n/2} \det(\Sigma_n^{-1})^{1/2} \exp(\Leftrightarrow \frac{1}{2} X_n^* \Sigma_n^{-1} X_n),$$

where Σ_n^{-1} is given by (10). Denoting the eigenvalues of Σ_n^{-1} by $\lambda_i, i = 0, \dots, n \Leftrightarrow 1$, we get the log likelihood

$$l_n(\alpha, \beta) = \Leftrightarrow \frac{n}{2} \log(2\pi) + \frac{1}{2} \sum_{i=0}^{n-1} \log(\lambda_i) \Leftrightarrow \frac{1}{2} X_n^* \Sigma_n^{-1} X_n.$$

The actual form of the eigenvalues is given in the Appendix. The maximum of $l_n(\alpha, \beta)$ can be found by standard numerical methods. Usually, it will be

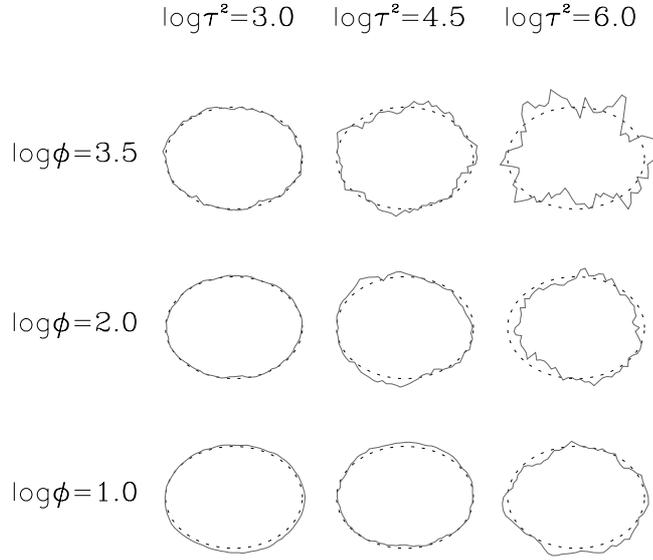


Figure 10: Simulations from the first-order Markov model for different values of ϕ and τ^2 . We have simulated from Model I, cf. Section 3.1, with \mathcal{C} taken to be an ellipse.

a good idea to estimate α and β for a collection of values of n and as a model check investigate whether the estimation is stable. Given that the model is suitable it appears to be a good idea to base the estimation on an n as large as possible.

Note that

$$\begin{aligned} X_n^* \Sigma_n^{-1} X_n &= \alpha \frac{1}{n} \sum_{i=0}^{n-1} X(t_i)^2 + \beta \frac{1}{n} \sum_{i=0}^{n-1} \left(\frac{X(t_i) \Leftrightarrow X(t_{i-1})}{1/n} \right)^2 \\ &= \alpha Y_n + \beta Z_n, \end{aligned}$$

say. The sufficient statistics (Y_n, Z_n) have a nice geometrical interpretation. Thus, Y_n is a discrete measure of the distance between the observed and the expected curve while Z_n compares local orientation,

$$Z_n = \frac{1}{n} \sum_{i=0}^{n-1} f(\Psi_i)^2,$$

where $\Psi_i = \Phi_{F_i} \Leftrightarrow \Phi_{c_i}$, and Φ_{F_i} and Φ_{c_i} are the angles that the line segments $F(t_i) \Leftrightarrow F(t_{i-1})$ and $c(t_i) \Leftrightarrow c(t_{i-1})$ make with a fixed axis, respectively. Fur-

thermore, $f = \tan$ or \sin , depending on whether $X(t)$ has been constructed using Model I or II of Section 3. Note also the connection between $X_n^* \Sigma_n^{-1} X_n$ and the sum of the external and internal energy from the theory of snakes, cf. (1).

4.2 Second-order Markov model

In Grenander (1993, p. 484) a model class with an approximate second-order Markov property is suggested. This class is characterized by the fact that $\{X(t)\}$ has mean zero, variance τ^2 and correlation function

$$\rho(h) = \frac{(\psi_2\psi_3 + \psi_1\psi_4)\psi_1(h)\psi_3(h) + (\psi_2\psi_3 \Leftrightarrow \psi_1\psi_4)\psi_2(h)\psi_4(h)}{\psi_1\psi_2 + \psi_3\psi_4}, \quad (11)$$

$$\psi > 0, \quad 0 \leq h \leq 1,$$

where we use the notation

$$\begin{aligned} \psi_1(h) &= \cos(\psi(h \Leftrightarrow 1/2)) & \psi_2(h) &= \sin(\psi(h \Leftrightarrow 1/2)) \\ \psi_3(h) &= \cosh(\psi(h \Leftrightarrow 1/2)) & \psi_4(h) &= \sinh(\psi(h \Leftrightarrow 1/2)) \\ \psi_i &= \psi_i(1), i = 1, \dots, 4. \end{aligned}$$

The explicit form of the correlation function is not given in Grenander (1993), but can be derived from the spectral density of $X(t)$, as demonstrated in the Appendix. Note that the correlation function only depends on ψ .

In Grenander (1993, p. 484), it is mentioned that the finite-dimensional distributions of $X(t)$ have a second-order Markov property, approximately. Recall that a multivariate normally distributed random vector

$$(X_0, \dots, X_{n-1}) \sim N_n(0, \Sigma)$$

has a second-order Markov property if it is Markov with respect to the second-order cyclic graph, which is given by

$$i \sim j \Leftrightarrow i = j \Leftrightarrow 2, j \Leftrightarrow 1, j + 1, j + 2 \pmod n.$$

The result by Grenander is formulated in the proposition below. For a $n \times n$ circular matrix $\{b_{ij}\}_{i,j=0}^{n-1}$ we use the notation $\text{circl}(a_0, a_1, a_2)$ if

$$b_{ij} = \begin{cases} a_0 & i = j \\ a_1 & i = j \Leftrightarrow 1, j + 1 \pmod n \\ a_2 & i = j \Leftrightarrow 2, j + 2 \pmod n \\ 0 & \text{otherwise.} \end{cases}$$

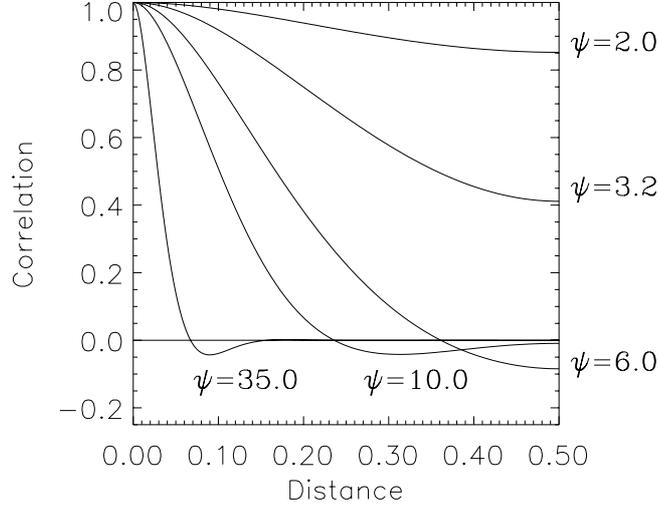


Figure 11: Correlation functions for the second-order Markov model for different values of ψ .

Proposition 4 *Let $\{X(t)\}$ be a cyclic and stationary Gaussian process with zero mean, variance τ^2 and correlation function (11). Furthermore, let $X^{(n)}(t)$ be constructed as in Proposition 3, but with*

$$\Sigma_n^{-1} = \text{circl}(\alpha/n + 6\gamma n^3, \Leftrightarrow 4\gamma n^3, \gamma n^3), \quad \alpha, \gamma > 0. \quad (12)$$

Then, $X^{(n)} = (X^{(n)}(t_0), \dots, X^{(n)}(t_{n-1}))$ is second-order Markov and $\{X^{(n)}(t)\}$ converges weakly to $\{X(t)\}$. Here, the 1-1 correspondence between (τ^2, ψ) and (α, γ) is

$$4\psi^4 = \alpha/\gamma \quad \tau^2 = \frac{\psi}{2\alpha} \frac{\psi_1\psi_2 + \psi_3\psi_4}{(\psi_2\psi_3)^2 + (\psi_1\psi_4)^2}.$$

The correlation structure of $X^{(n)}$, from which the second-order Markov property follows, is given by

$$\begin{aligned} & \rho(X^{(n)}(t_i), X^{(n)}(t_j) | X^{(n)} \setminus \{X^{(n)}(t_i), X^{(n)}(t_j)\}) \\ &= \begin{cases} 1 & i = j \\ \frac{1}{(\psi/n)^4 + 3/2} & i = j \Leftrightarrow 1, j + 1 \pmod n \\ \frac{1/4}{(\psi/n)^4 + 3/2} & i = j \Leftrightarrow 2, j + 2 \pmod n \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

As in the first order case the parameter τ^2 is a measure of the overall difference, while ψ regulates the smoothness of curve, cf. Figure 12.

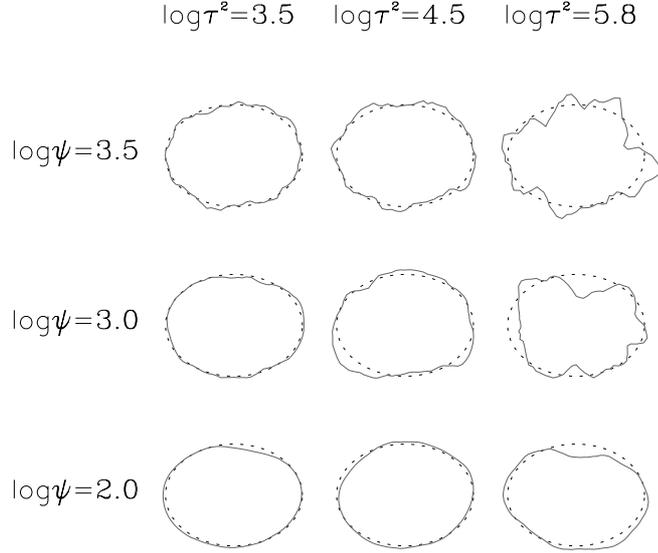


Figure 12: Simulations from the second-order Markov model for different values of ψ and τ^2 . We have simulated from Model I, cf. Section 3.1, with \mathcal{C} equal to an ellipse.

Similarly to the first-order case we can estimate the parameters by an approximate likelihood analysis. In this case

$$\begin{aligned} X_n^* \Sigma_n^{-1} X_n &= \alpha \frac{1}{n} \sum_{i=0}^{n-1} X(t_i)^2 + \gamma \frac{1}{n} \sum_{i=0}^{n-1} \left(\frac{X(t_i) \Leftrightarrow 2X(t_{i-1}) + X(t_{i-2})}{(1/n)^2} \right)^2 \\ &= \alpha Y_n + \gamma V_n. \end{aligned}$$

V_n can be interpreted geometrically as a discrete measure of the change in local orientation

$$V_n = \frac{1}{n} \sum_{i=0}^{n-1} \left(\frac{f(\Psi_i) \Leftrightarrow f(\Psi_{i-1})}{(1/n)^2} \right)^2,$$

where f and Ψ_i are defined as in the first-order case.

5 The application

For each nuclear profile we calculated the boundary length T and considered the $n = 50$ points

$$(F(Tt_0), \dots, F(Tt_{n-1})), \quad t_i = i/n, \quad i = 0, \dots, n \Leftrightarrow 1,$$

collected (approximately) equidistantly on the boundary of the nuclear profile. For \mathcal{C} we chose the ellipse from Section 2. This is the set-up in Model II, cf. Section 3.2. The reason for preferring Model II to Model I was that Model II was expected to be a more sensitive tool for distinguishing between the profiles from the malignant and the benign tumour.

Now consider the corresponding n observations from the normalized residual process

$$X_n = (X(Tt_0), \dots, X(Tt_{n-1}))/T, \quad t_i = i/n, \quad i = 0, \dots, n \Leftrightarrow 1.$$

In Section 4, we have presented two models for the normalized residual process. Our initial model was a general second-order Markov model

$$M_0 : X_n \sim N_n(\mathbf{0}, \Sigma_n),$$

where

$$\Sigma_n^{-1} = \text{circ1}(\alpha/n + 2\beta n + 6\gamma n^3, \Leftrightarrow \beta n \Leftrightarrow 4\gamma n^3, \gamma n^3), \quad \alpha > 0, \quad \beta, \gamma \geq 0,$$

see also the Appendix. For all the profiles, we estimated the covariance matrix, using the ellipses fitted in Section 2. For some of the profiles, the ellipse and the covariance were also estimated simultaneously. The ellipses fitted in this way did only differ slightly from those determined in Section 2.

To see whether the first- or second-order Markov model from Section 4 could be used we investigated the two hypotheses

$$H_1 : \gamma = 0 \quad H_2 : \beta = 0,$$

with corresponding models M_1 and M_2 . Let L_i be the maximized likelihood under the model M_i . We then tested whether $M_i, i = 1, 2$, was a reasonable simplification of M_0 by performing the usual likelihood ratio test

$$\Leftrightarrow 2 \log(L_i/L_0) \approx \chi^2(1).$$

The test probabilities for H_1 lay in the range 0%-1%, and for H_2 in the range 95%-100%. Hence the second-order Markov model with $\beta = 0$ described in Section 4.2 was the most appropriate. Since the nuclear profiles look like the ones simulated in Figure 12 this was of course not surprising.

Figure 13 shows the estimators of (α, γ) under the model M_2 . Note that it is possible to distinguish between the two samples.

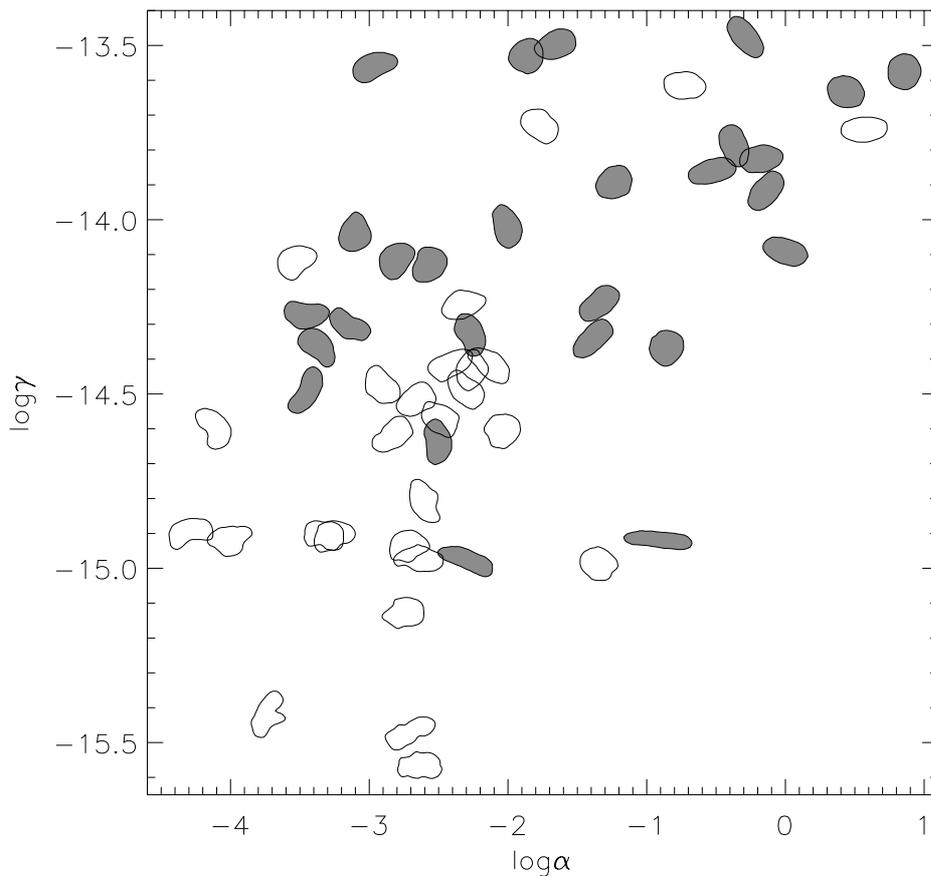


Figure 13: The estimators of (α, γ) under the model M_2 . The white nuclei are from the malignant tumour while the black nuclei are from the benign tumour.

To verify that the estimators do not depend on the number n of landmarks, we fitted the model for landmark numbers between 30 and 90. From Figure 14 it is clear, that the estimation is stable.

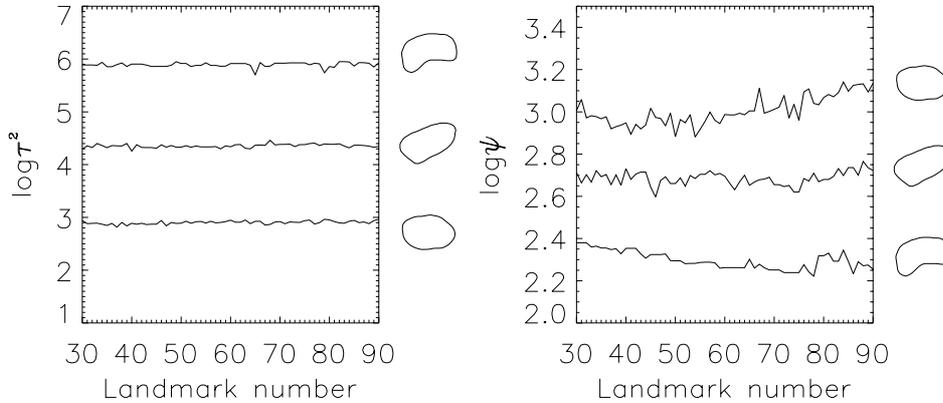


Figure 14: The estimators of $\log \tau^2$ and $\log \psi$ for different numbers of landmarks.

6 Discussion

The smoothness of the sample paths of the residual process is an important topic when constructing realistic models. It follows from (16) in the Appendix below that the processes considered in Sections 4.1 and 4.2 have a spectral density of the form

$$(\alpha + \beta(2\pi s)^2)^{-1} \quad s = 0, 1, \dots \quad (13)$$

$$(\alpha + \gamma(2\pi s)^4)^{-1} \quad s = 0, 1, \dots, \quad (14)$$

respectively. Since (13) decreases as $1/s^2$ and (14) decreases as $1/s^4$ the sample paths of the first type of process are continuous, while the sample paths of the second type are continuously differentiable, cf. Cramer and Leadbetter (1967, p. 181). For each nuclear profile it seems natural to assume that the observed curve \mathcal{F} is differentiable, and since the expected curve \mathcal{C} is infinitely often differentiable the model from Section 4.2 appear to be the most appropriate in our application, also from the point of view of smoothness.

The model presented in 2 dimensions has a natural extension to 3 dimensions. An observed surface $\mathcal{F} = \{F(t) : t \in \Omega \subseteq \mathbb{R}^2\}$ is again assumed to be a stochastic deformation of an expected surface $\mathcal{C} = \{c(t) : t \in \Omega \subseteq \mathbb{R}^2\}$, so that

$$F(t) = c(t) + X(t)\omega(t), \quad t \in \Omega \subseteq \mathbb{R}^2.$$

As in 2 dimensions $\omega(t)$ represents an outer unit normal vector to \mathcal{C} at $c(t)$. Suppose we observe

$$X_n = (X(t_0), \dots, X(t_{n-1})), \quad t_i \in \Omega, \quad i = 0, \dots, n \Leftrightarrow 1.$$

A natural model would then be to let X_n have a multivariate distribution with zero mean and a covariance matrix with entries

$$\text{cov}(X(t_i), X(t_j)) = \tau^2 \rho(d_{\mathcal{C}}(t_i, t_j)).$$

Here, $d_{\mathcal{C}}(t_i, t_j)$ is the distance on the surface between $c(t_i)$ and $c(t_j)$. The difficult part is to find an appropriate correlation function. In this paper we have considered two possibilities as described in Section 4.1 and Section 4.2.

Another and much more simple approach would be to remain in 2 dimensions by placing a random plane through the center of the nucleus. An analysis similar to the one described in the present paper could then be carried out.

Consider again the normalized residual process $\{X_1(t) : 0 \leq t \leq 1\}$. Instead of having zero mean a mean value of the residual process given by either

$$\mu_1(t) = \alpha_1 \cos(6\pi t + \omega_1), \quad 0 \leq t \leq 1,$$

or

$$\mu_2(t) = \alpha_2 \cos(8\pi t + \omega_2), \quad 0 \leq t \leq 1,$$

or a sum of $\mu_1(t)$ and $\mu_2(t)$ would allow for a systematic variation from the shape of an ellipse. Still modelling the correlation structure by the second-order Markov property it might be even more easy to distinguish between the two samples.

Acknowledgement

The authors thank Hans Jørgen Gundersen, Flemming Sørensen and colleagues from Department of Mathematical Sciences, University of Aarhus, for fruitful discussions.

Appendix

In this appendix we will derive the explicit form (11) of the correlation function for the second-order model. We will use an approach such that we at the same time derive $\rho(h)$ for the first-order model, cf. (6).

Thus, consider the matrix

$$\Sigma_n^{-1} = \text{circl}(\alpha/n + 2\beta n + 6\gamma n^3, \Leftrightarrow \beta n \Leftrightarrow 4\gamma n^3, \gamma n^3), \quad \alpha > 0, \quad \beta, \gamma \geq 0.$$

For $v = (v_0, \dots, v_{n-1}) \in \mathbb{R}^n$ we have that

$$v \Sigma_n^{-1} v^* = \alpha/n \sum_{i=0}^{n-1} v_i^2 + \beta n \sum_{i=0}^{n-1} (v_i \Leftrightarrow v_{i-1})^2 + \gamma n^3 \sum_{i=0}^{n-1} (v_i \Leftrightarrow 2v_{i-1} + v_{i-2})^2,$$

where $v_{-1} = v_{n-1}$ and $v_{-2} = v_{n-2}$. Therefore, Σ_n^{-1} is positive definite. Note that $\gamma = 0$ is the case treated in Proposition 3 and $\beta = 0$ is the case treated in Proposition 4.

Since Σ_n^{-1} is circular we have that Σ_n is also circular. The eigenvalues of Σ_n^{-1} are given by, cf. Anderson (1958, p. 280, 282),

$$\lambda_s = (\alpha/n + 2\beta n + 6\gamma n^3) \Leftrightarrow (\beta n + 4\gamma n^3) 2 \cos(2\pi s/n) + \gamma n^3 2 \cos(4\pi s/n),$$

with corresponding orthonormal (complex) eigenvectors $u_s, s = 0, \dots, n \Leftrightarrow 1$, where

$$(u_s)_l = e^{2\pi l s i/n} / \sqrt{n}, \quad l = 0, \dots, n \Leftrightarrow 1.$$

Note that $u_0 = (1, \dots, 1) / \sqrt{n}$. Let

$$\Lambda = \text{diag}(\lambda_0, \dots, \lambda_{n-1}) \text{ and } U = \begin{pmatrix} u_0 \\ \vdots \\ u_{n-1} \end{pmatrix}.$$

Then, we have for $l = 0, 1, \dots, n \Leftrightarrow 1$ that

$$\begin{aligned} (\Sigma_n)_{l,0} &= (U \Lambda^{-1} U^*)_{l,0} = \sum_{s=0}^{n-1} \frac{1}{\sqrt{n}} (U \Lambda^{-1})_{l,s} \\ &= \sum_{s=0}^{n-1} \frac{e^{2\pi l s i/n}}{n \lambda_s} = \sum_{s=0}^{n-1} \frac{\cos(2\pi l s/n)}{n \lambda_s}. \end{aligned} \tag{15}$$

Now, we will study a sequence $X^{(n)}(t)$ of cyclic Gaussian processes, constructed as in Propositions 3 and 4. Thus, let

$$(X_{n0}, \dots, X_{n(n-1)}) \sim N_n(0, \Sigma_n)$$

and define a continuous process $\{X^{(n)}(t) : 0 \leq t \leq 1\}$ by

$$X^{(n)}(j/n) = X_{nj} \quad j = 0, \dots, n \Leftrightarrow 1,$$

with linear interpolation in between, and with $X^{(n)}(0) = X^{(n)}(1)$. Using (15), we find the following expression for the covariance between $X^{(n)}(0)$ and $X^{(n)}([nh]/n)$, $0 \leq h < 1$,

$$\text{Cov}(X^{(n)}(0), X^{(n)}([nh]/n)) = \text{Cov}(X_{n0}, X_{n([nh])}) = \sum_{s=0}^{n-1} \frac{\cos(2\pi[nh]s/n)}{n\lambda_s}.$$

Assuming that at least one of the parameters β and γ is positive, we obtain, using dominated convergence,

$$\lim_{n \rightarrow \infty} \text{Cov}(X^{(n)}(0), X^{(n)}([nh]/n)) = \frac{1}{\alpha} + 2 \sum_{s=1}^{\infty} \frac{\cos(2\pi hs)}{\alpha + \beta(2\pi s)^2 + \gamma(2\pi s)^4}. \quad (16)$$

By construction $X^{(n)}(t)$, $0 \leq t < 1$, is given by

$$X^{(n)}(t) = X_{n([nt])} + n(t \Leftrightarrow [nt]/n)(X_{n([nt]+1)} \Leftrightarrow X_{n([nt])}).$$

Hence we have

$$\begin{aligned} |X^{(n)}(h) \Leftrightarrow X_{n([nh])}| &= n(h \Leftrightarrow [nh]/n) |X_{n([nh]+1)} \Leftrightarrow X_{n([nh])}| \\ &\leq |X_{n([nh]+1)} \Leftrightarrow X_{n([nh])}| \end{aligned}$$

and

$$\begin{aligned} \text{E}((X_{n([nh]+1)} \Leftrightarrow X_{n([nh])})^2) &= \text{E}((X_{n1} \Leftrightarrow X_{n0})^2) \\ &= 2(\text{Var}X_{n0} \Leftrightarrow \text{Cov}(X_{n1}, X_{n0})) \\ &= 2\left(\sum_{s=0}^{n-1} \frac{1}{n\lambda_s} \Leftrightarrow \sum_{s=0}^{n-1} \frac{\cos(2\pi s/n)}{n\lambda_s}\right) \\ &\rightarrow 0 \text{ for } n \rightarrow \infty. \end{aligned} \quad (17)$$

From (16) and (17) it is clear that the limiting covariance function is given by the Fourier series

$$\sigma(h) = \frac{1}{\alpha} + 2 \sum_{s=1}^{\infty} \frac{\cos(2\pi hs)}{\alpha + \beta(2\pi s)^2 + \gamma(2\pi s)^4}, \quad 0 \leq h \leq 1.$$

In order to find an explicit expression of $\sigma(h)$ note that in the sense of generalized functions and the theory of distributions

$$\begin{aligned} \alpha\sigma(h) \Leftrightarrow \beta\sigma''(h) + \gamma\sigma''''(h) &= 1 + 2 \sum_{s=1}^{\infty} \cos(2\pi hs) \\ &= \sum_{s=-\infty}^{\infty} \cos(2\pi hs) = \sum_{s=-\infty}^{\infty} \delta_s, \end{aligned}$$

where δ_s denotes the δ -function at s . Therefore, $\sigma(h)$ is the solution of a homogeneous linear differential equation with constant coefficients, cf. Hirsch and Smale (1974, p. 138). In general, $\sigma(h)$ is thus a linear combination of exponential and trigonometric functions depending on the roots of the characteristic polynomial

$$\alpha \Leftrightarrow \beta y^2 + \gamma y^4 = 0.$$

We will for simplicity only consider the two cases previously mentioned, $\beta > 0, \gamma = 0$ and $\beta = 0, \gamma > 0$.

Before treating these two cases, it is convenient to change the interval from $[0, 1]$ to $[\Leftrightarrow 1/2, 1/2]$. Thus, let

$$\sigma_*(h) = \sigma(h + 1/2), \quad \Leftrightarrow 1/2 \leq h \leq 1/2.$$

The function $\sigma_*(h)$ is the solution to the same homogeneous linear differential equation,

$$\alpha\sigma_*(h) \Leftrightarrow \beta\sigma_*''(h) + \gamma\sigma_*''''(h) = 0. \quad (18)$$

First consider the case $\beta > 0, \gamma = 0$. With $\phi^2 = \alpha/\beta$ the solution of (18) is given by, cf. Hirsch and Smale (1974, p. 139),

$$\sigma_*(h) = c_1 e^{\phi h} + c_2 e^{-\phi h}, \quad \Leftrightarrow 1/2 \leq h \leq 1/2.$$

From $\sigma(h) = \sigma(1 \Leftrightarrow h), 0 \leq h \leq 1$, it follows that

$$\sigma_*(h) = \sigma_*(\Leftrightarrow h), \quad \Leftrightarrow 1/2 \leq h \leq 1/2,$$

and hence we have $c_1 = c_2 = c$. Using that the first Fourier coefficient is given by $1/\alpha$ we can find c from the equation

$$1/\alpha = 2 \int_0^{1/2} \sigma_*(h) dh, \quad (19)$$

and we obtain the form (6) of $\rho(h)$, stated in the main text, and the correspondence between (τ^2, ϕ) and (α, γ) given in Proposition 3.

Now consider the case $\beta = 0, \gamma > 0$. Again, using $\sigma_*(h) = \sigma_*(\Leftrightarrow h)$ and letting $4\psi^4 = \alpha/\gamma$, the solution of (18) is given by, cf. Hirsch and Smale (1974, p. 139)

$$\sigma_*(h) = c_1(e^{\psi h} + e^{-\psi h}) \cos(\psi h) + c_2(e^{\psi h} \Leftrightarrow e^{-\psi h}) \sin(\psi h), \quad \Leftrightarrow 1/2 \leq h \leq 1/2.$$

The constants c_1 and c_2 can be found from (19) and $\sigma'_*(1/2) = 0$, and we obtain the explicit form (11) of $\rho(h)$ and the correspondence between (τ^2, ψ) and (α, γ) given in Proposition 4.

References

- Anderson, T.W. (1958). *The Statistical Analysis of Time Series*. John Wiley and Sons, New York.
- Cramer, H and Leadbetter, M.R. (1967). *Stationary and Related Stochastic Processes*. Wiley, New York.
- Grenander, U. (1993). *General Pattern Theory*. Oxford University Press, Oxford.
- Grenander, U. and Manbeck, K.M. (1993). A stochastic model for defect detection in potatoes. *J. Comp. Graph. Statist.* **2**, 131-151.
- Hirsch, M.W. and Smale, S. (1974). *Differential Equations, Dynamical Systems and Linear Algebra*. Academic Press, San Diego.
- Jensen, E.B.V. (1998). *Local Stereology*. World Scientific, Singapore.
- Jensen, E.B. and Sørensen, F.B. (1991). A note on stereological estimation of the volume-weighted second moment of particle volume. *J. Microsc.* **164**, 21-27.
- Kass, M., Witkin, A., Terzopoulos, D. (1988). Snakes: Active Contour Models. *International Journal of Computer Vision* **1**, 321-331.

Kent, J.T., Mardia, K.V. and Walder, A.N. (1996). Conditional cyclic Markov random fields. *Adv. Appl. Probab. (SGSA)* **28**, 1-12.

Lauritzen, S.L. (1996). *Graphical Models*. Oxford University Press, Oxford.

Rue, H. and Syversveen, A.R. (1998). Bayesian object recognition with Baddeley's Delta Loss. *Adv. Appl. Prob. (SGSA)* **30**, 64-84.

Sørensen, F.B. (1991). Stereological estimation of mean and variance of nuclear volume from vertical sections. *J. Microsc.* **162**, 203-229.