Chapter 2

Ridge Definitions

2.1 Ridge Definitions

Numerous attempts have been made to construct ridges for use in image and shape analysis. Consequently many definitions for ridges can be found in the literature. Each definition has its advantages and disadvantages, but it is desirable that ridges satisfy certain properties. Firstly, the process which identifies ridges should be *local*. That is, a ridge point should be determined solely by the information in a local neighborhood of the point. Secondly, the ridges should be *invariant* with respect to the following transformations:

- translations in the spatial variables,
- rotations in the spatial variables,
- uniform magnification in the spatial variables, and
- monotonic transformations of the intensity function.

The first two invariances simply mean that the operations of ridge construction and rigid motions should commute. The ridges of a rotated/translated object should be the rotated/translated ridges of the original object. The third invariance means that the ridge construction should be independent of the units of measurement, since a change of units is equivalent to a uniform magnification in the spatial variables. The ridges of an object which is doubled in size should be the ridges, doubled in size, of the original object. The fourth property is desired since the data, as sampled by the imaging device, may be modified in a monotonic way to meet the requirements of a display device. For example, if 12–bit data are to be displayed on an 8–bit display, the data are typically transformed using a monotonic affine transformation. We would like the ridge structures not to be affected by this transformation. A study of invariance and its applications to computer vision can be found in ter Haar Romeny, Florak, Koenderink & Viergever (1991).

The definition for ridges as *slope district boundaries* is used in Colchester (1990) and Griffin, Colchester & Robinson (1991) for segmenting intensity images. The graph of the intensity values is searched for maxima, minima, and saddle points. Slope lines, which are integral curves of the intensity gradient, are drawn from minima to saddles (course lines) and from saddles to maxima (ridge lines). The regions enclosed by the slope lines are called slope districts and are the primitive regions of the segmentation.

The slope district construction has also been used on hypersurfaces. The Blum medial construction for 3–dimensional objects (Blum 1973, Blum & Nagel 1978) yields a medial surface which encodes the shape of the object. Each surface point has an associated radius for the maximal sphere centered at the point and contained entirely inside the object. The medial surface is typically as complicated as the object itself. Attempts have been made to simplify this surface by segmenting it using slope districts (Nackman 1982, Nackman 1984). In this case the domain of the radius function is the medial surface, as compared to the planar domain for intensity images. The graph of the radius function is searched for maxima, minima, and saddle points. As before, ridge and course lines are drawn connecting the critical points.

The ridges from the slope district definition are clearly invariant under spatial translations, spatial rotations, and uniform spatial magnifications. They are also invariant under monotonic transformations of intensity since the gradient vector fields of the intensity and the transformed intensity have the same directions (but not necessarily the same lengths).

An undesirable consequence of the slope district construction is that the ridge and course lines are determined by a nonlocal process. Consider a ridge line whose endpoints are a maximum and a saddle. Each ridge point occurs simply because it is on the flow line connecting the maximum and saddle. No local geometric information at the point is used to establish its identity as a ridge. If the intensity function is perturbed in a small neighborhood of the saddle, the effect of the perturbation is propagated to *all* the ridge points on the flow to the maximum, even if the intensity values near the maximum have not been perturbed at all. Thus, the ridge points at the maximum have been relocated even though no change in intensity has occurred in that region. Another deficiency in the slope district construction is that it may ignore small ridges which lie on the flank of a larger ridge. The small ridges are frequently not formed by the presence of local extrema of the intensity.

A discrete definition for ridges appears in Crowley & Parker (1984), where the underlying function is the intensity convolved with a difference of low–pass transforms. Given a discrete 2–dimensional image, ridges are defined as those pixels for which the function has a positive local maximum in one of the 4 directions associated with the 8–neighborhood of the pixel. The generated ridges are invariant under spatial translations and monotonic transformations of the function. The ridges are not preserved by rotations. At any point where the function has a local maximum with respect to some direction V, that point will be labeled as a ridge in an orientation which aligns V with one of the 4 special directions, even if the point was not labeled as a ridge in its original orientation.

A more successful definition for ridges is called the *height definition*. A history of the attempt to define ridges of this type, dating back to 1852, is found in Koenderink & van Doorn (1993). More recent work can be found in Haralick (1983) and Morse, Pizer & Liu (1993). The definition is based on computing local maxima for the intensity (or height) function in special directions. The ridge construction by this definition is a local process. Moreover, the ridges are invariant under spatial translations, spatial rotations, uniform spatial magnifications, but not under monotonic transformations of the height function. I discuss the height definition in Section 2.3.

Other researchers have considered defining ridges using the ideas of differential geometry. The definitions are local and involve measurements of curvatures associated with graphs or surfaces in general. I will consider two such definitions.

The first of the differential geometric definitions, called the *principal direction definition*, applies to n–dimensional hypersurfaces. The standard analysis of hypersurfaces uses a parameterization to represent the hypersurface. For example, if the hypersurface is the graph for a function $f(x)$, the parameterization is $(x, f(x))$. Some computational vision applications involve hypersurfaces for which a natural parameterization is not immediately apparent. If the hypersurface is implicitly defined as the level set for a function, a definition can still be given for ridges on the hypersurface. Every hypersurface (for example, graphs of functions defined on $\mathbb{R}^n$) can be described this way, so there is no not loss of generality. Whereas other ridge definitions require a spatial parameterization of the graph, the principal direction definition does not require one to define ridges.

The ridges are constructed as local extrema of principal curvatures, where the differentiation is taken in principal directions. A geometrically intuitive discussion of this type ridge is found in Koenderink (1991); a formal mathematical analysis is given in Bucchi (1991). The ridge construction by the principal direction definition is a local process. For a parameterized hypersurface, the ridges are invariant under diffeomorphisms on the parameter space. Also, the ridges are invariant under spatial translations, spatial rotations, and uniform spatial magnifications, where the transformations are applied to the full space $\mathbb{R}^{n+1}$ which contains the hypersurface. In regard to graphs of functions with domain $\mathbb{R}^n$, the ridges are not invariant under monotonic changes of the defining function. An application of the principal direction definition can be found in Thirion & Gourdon (1992). Surfaces representing a skull were obtained from 3–dimensional data sets. Ridges on the surfaces were constructed and used as landmarks for image registration. The principal direction definition is discussed in Section 2.4.

The second of the differential geometric definitions, called the *level definition*, considers curvature properties of level sets for smooth functions. The constraint of smoothness for intensity functions is consistent with the way that a scene is observed using finite width apertures (Florack, ter Haar Romeny, Koenderink & Viergever 1993). A motivating example is the function $f(x, y) = 1 - (x/a)^2 - (y/b)^2$ where $a > b$, where for illustration suppose a is very much larger than b. Treating the graph as mountain terrain, a person walking at constant altitude around the mountain might label the ridges as those points where the change in his direction of walking changes the most. In this case, the ridge points are identified at places along the path where $y = 0$. The paths of constant altitude are the level curves of the function. The places where the direction (level curve tangent) changes most rapidly are those points for which the level curve curvature is locally optimal. The ridge construction is local, and the ridges are invariant under spatial translations, spatial rotations, uniform spatial magnifications, and monotonic transformations of the intensity function. Applications of the level definition to computer vision problems are found in Gauch, Oliver & Pizer (1988). I discuss the level definition in Section 2.5. It turns out to be exactly the principal direction definition, but applied to level surfaces of the graph instead of the entire graph itself.

Finally, I give a ridge definition, called the *nonmetric definition*, which is similar to the principal direction definition, but applies only to graphs of functions. The ridges produced by this definition are often qualitatively similar to those produced by the principal direction definition. However, unlike the principal direction definition, this ridge construction does not use the metric of the graph. I discuss the nonmetric definition in Section 2.6.

Section 2.7 contains the application of the height, principal direction, level, and nonmetric definitions of ridges to the MR image of Figure 1.4. A discussion of the findings for the different ridge definitions is given in Section 2.8.

The wide range of results on ridges in image analysis described above clearly indicates that numerous authors have found a need for ridges in image analysis as a means of describing object shape. The purpose of this chapter is to provide a formal mathematical setting for ridge definitions, to generalize the ideas to higher dimensions, and to analyze the properties of the various definitions. Although the emphasis of the development is on ridges, the constructions naturally include the counterparts of ridges, namely *valleys*. For completeness, I include definitions for both ridge and valley points. Collectively all such points will be called *crease points*. They are located on the graph of f and therefore have the form $(x, f(x))$; however, I will also refer to x as a crease point. This identification should not create any confusion in the development.

2.2 Directional Derivatives

Throughout the chapter I use the following notations. The partial derivatives of a function $f : \mathbb{R}^n \to \mathbb{R}$ are denoted by subscripting f with the appropriate variable. If $x = (x_1, \ldots, x_n)$, then f_{x_i} is the first partial derivative of f with respect to x_i. The *gradient* of f is the vector

$\nabla f = (f_{x_1}, \ldots, f_{x_n})$. The matrix of second partial derivatives of f, called the *Hessian* of f, is denoted by $\mathrm{H}(f) = [f_{x_i x_j}]$. The function f is said to be a C^k function if its partial derivatives through order k are continuous.

All crease definitions require finding local extrema of functions in special directions. *Directional derivatives* measure how a function varies when restricted to a subset of its domain. The zeros of first–order directional derivatives will be the candidate crease points. A candidate crease point will be classified as a ridge point or valley point depending on the information obtained from second–order directional derivatives. I present two definitions for each order directional derivative. The first definition assumes that the direction vectors are constant, but the second definition allows the vectors to be variable. For first–order directional derivatives, the two definitions are equivalent. However, second–order directional derivatives are computed as iterations of differential operators which depend on the direction vectors, so the two definitions produce different results.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2 function. Let $v = \in \mathbb{R}^n$ be a constant nonzero vector. The *first derivative of f at x in the direction v* is defined by

$$(v \cdot \nabla) f(x) = \sum_{i=1}^{n} v_i \frac{\partial f(x)}{\partial x_i},$$

where '$\cdot$' indicates the dot product of vectors. The definition is motivated by calculus of a function of a single variable. At a point $p \in \mathbb{R}^n$, let $\xi(t)$ be a differentiable curve such that $\xi(0) = x$ and $\xi'(0) = v$. Define the function $\phi(t) = f(\xi(t))$ and compute the ordinary derivative with respect to t. An application of the chain rule yields $\phi'(t) = \xi'(t) \cdot \nabla f(\xi(t))$. At $t = 0$, $\phi'(0) = v \cdot \nabla f(x)$. Note that the first directional derivative is independent of the chosen path as long as the tangent vector at $t = 0$ is v.

Now let the directions be dependent on the points x at which the derivative measurement is made. Let $v(x)$ be a nonzero vector. The *first derivative of f at x in the direction $v(x)$* is defined by

$$D_v f(x) = \sum_{i=1}^{n} v_i(x) \frac{\partial f(x)}{\partial x_i}.$$

At a particular point y, let $u = v(y)$; then $(u \cdot \nabla) f(y) = D_v f(y)$, so the definitions are equivalent.

Second derivatives in specified directions can be similarly defined. Let v be a constant nonzero vector. The *second derivative of f at x in the direction v* is defined by

$$(v \cdot \nabla)^2 f(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} v_i \frac{\partial^2 f(x)}{\partial x_i \partial x_j} v_j = v^t \mathrm{H}(f) v.$$

If $\xi(t) = x + tv$ and $\phi(t) = f(\xi(t))$, then $\phi(0) = f(x)$, $\phi'(0) = v \cdot \nabla f(x)$, and $\phi''(0) = v^t \mathrm{H}(f) v$. Therefore, $(v \cdot \nabla)^2 f(x)$ measures the second derivative along a *linear* path through x in the direction v. Mixed directional derivatives are also allowed. Let u and v be constant nonzero vectors. The *second derivative of f at x in the directions u and v* is defined by

$$(v \cdot \nabla)(u \cdot \nabla) f(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} u_i \frac{\partial^2 f(x)}{\partial x_i \partial x_j} v_j = u^t \mathrm{H}(f) v.$$

Since f is C^2, the order of differentiation is irrelevant (the matrix $\mathrm{H}(f)$ is symmetric), so $(v \cdot \nabla)(u \cdot \nabla) f(x) = (u \cdot \nabla)(v \cdot \nabla) f(x)$. Now define $\xi(s,t) = x + su + tv$ and $\phi(s,t) = f(\xi(s,t))$; then $\phi(0,0) = f(x)$, $\phi_s(0,0) = u \cdot \nabla f(x)$, $\phi_t(0,0) = v \cdot \nabla f(x)$, and $\phi_{st}(0,0) = u^t \mathrm{H}(f(x)) v = v^t \mathrm{H}(f(x)) u = \phi_{ts}(0,0)$.

The second derivative definition using constant direction vectors is dependent on the parameterization $\xi(s,t) = x + su + tv$. Other parameterizations may lead to different results. If *variable* nonzero vector fields $u(x)$ and $v(x)$ are specified and second derivatives are computed, there are some problems to consider. The parameterization $\xi(s,t)$ must have the properties $\xi(0,0) = x$, $\xi_s(s,t) = u(\xi(s,t))$, and $\xi_t(s,t) = v(\xi(s,t))$. For a smooth solution $\xi(s,t)$ to exist, it is necessary that $\xi_{st} = \xi_{ts}$, which implies that

$$0 = \frac{\partial}{\partial t} u(\xi(s,t)) - \frac{\partial}{\partial s} v(\xi(s,t)) = \frac{du(\xi)}{dx} v(\xi) - \frac{dv(\xi)}{dx} u(\xi) = [u,v](\xi),$$

where du/dx is the $n \times n$ matrix whose i^{th} row contains the partial derivatives of component u_i, and where $[u,v]$ is the Lie product of vector fields. Therefore, at every x it is necessary that the vector fields commute in the sense of Lie products: $[u,v](x) = 0$ for all x. If the vector fields commute, the *second directional derivative of f at x in the directions $u(x)$ and $v(x)$* is defined by

$$\begin{aligned} D_u D_v f(x) &= u(x)^t \mathrm{H}(f(x)) v(x) + \nabla f(x)^t \tfrac{du}{dx} v(x) \\ &= v(x)^t \mathrm{H}(f(x)) u(x) + \nabla f(x)^t \tfrac{dv}{dx} u(x) \\ &= D_v D_u f(x). \end{aligned}$$

Note that this definition reduces to the previous one when u and v are constant vectors.

The directional derivative definitions using constant vectors are what the height definition for creases uses; the quadratic form involving the Hessian will be used for second derivative tests. But in the principal direction definition variable vectors are used, so the other directional derivative definitions will be used in this case.

2.3 Height Definition

The height definition for ridges and valleys is based on a generalization of local extrema for real–valued functions of a vector variable. I give a brief summary of the construction of local extrema; the construction can be found in standard calculus books (Fulks 1978).

2.3.1 Extreme Points

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2 function. A point $x_0 \in \mathbb{R}^n$ is said to be a *critical point* for f if $(v \cdot \nabla) f(x_0) = 0$ for every direction v, which is equivalent to $\nabla f(x_0) = 0$. Critical points are classified via the following. The function f has

- a minimum at x_0 if $(v \cdot \nabla)^2 f(x_0) > 0$ for every direction v;
- a maximum at x_0 if $(v \cdot \nabla)^2 f(x_0) < 0$ for every direction v.

Any such point x_0 is called an *extreme point* and the corresponding function value is called an *extreme value*. It is only necessary to perform the tests for n linearly independent directions. Thus, x_0 is a critical point if $\nabla f(x_0) = 0$, and $f(x_0)$ is

- a minimum at x_0 if $\mathrm{H}(f(x_0))$ is positive definite (all eigenvalues are positive);
- a maximum at x_0 if $\mathrm{H}(f(x_0))$ is negative definite (all eigenvalues are negative).

2.3.2 Relative Extreme Points

I consider a more general definition for extreme points f. Let $v_1, \ldots, v_d$ be a set of constant linearly independent vectors in $\mathbb{R}^n$, where $1 \leq d \leq n$. When $d = n$, the vectors can be thought of as the columns of an $n \times n$ invertible matrix V. The matrices $\mathrm{H}(f)$ and $V^t \mathrm{H}(f) V$ have the same definiteness by Sylvester's Theorem (Horn & Johnson 1991). Therefore, the construction of local extrema in the previous subsection is equivalent to the following. The function f has

- a *minimum* at x_0 if $V^t \nabla f(x_0) = 0$ and $V^t \mathrm{H}(f(x_0))V$ is positive definite;
- a *maximum* at x_0 if $V^t \nabla f(x_0) = 0$ and $V^t \mathrm{H}(f(x_0))V$ is negative definite.

Now consider the cases $d < n$. Let V be the $n \times d$ matrix whose columns are the vectors $v_1, \ldots, v_d$. A search is made for local extrema of f restricted to the subspace spanned by the v_i. The function has

- a *relative minimum of type* $n - d$ at x_0 if $V^t \nabla f(x_0) = 0$ and $V^t \mathrm{H}(f(x_0))V$ is positive definite;
- a *relative maximum of type* $n - d$ at x_0 if $V^t \nabla f(x_0) = 0$ and $V^t \mathrm{H}(f(x_0))V$ is negative definite.

Such points x_0 are called *relative extreme points of type* $n - d$ *for* f *with respect to* V. The classification has the same form as the case $d = n$, but the $x_0 \in \mathbb{R}^n$ are now solutions to d equations in n unknowns. The solution sets are usually $(n - d)$–dimensional manifolds, hence the use of "type $n - d$" in the definition. Ridges and valleys will be defined as relative extreme points with respect to eigenvectors of the Hessian of f.

2.3.3 Crease Definitions

Define $W = -\mathrm{H}(f)$ and let κ_i and v_i, $1 \le i \le n$, be its eigenvalues and eigenvectors. A positive (negative) eigenvalue corresponds to convexity (concavity) of the graph of f in the corresponding eigendirection. Assume that the eigenvalues are ordered as $\kappa_1 \ge \cdots \ge \kappa_n$. In the definition assume that $1 \le d \le n$.

- A point x is a *ridge point of type* $n - d$ if $\kappa_d(x) > 0$ and x is a relative maximum point of type $n - d$ for f with respect to $V = [v_1 \cdots v_d]$. Since $-V^t \mathrm{H}(f)V = \mathrm{diag}\{\kappa_1 |v_1|^2, \ldots, \kappa_d |v_d|^2\}$ and since the eigenvalues are ordered, the test for a ridge point reduces to $V^t \nabla f(x) = 0$ and $\kappa_d(x) > 0$.

- A point x is a *valley point of type $n-d$* if $\kappa_{n-d+1}(x) < 0$ and x is a relative minimum point of type $n-d$ for f with respect to $V = [v_{n-d+1} \cdots v_n]$. Since $-V^t \mathrm{H}(f) V = \mathrm{diag}\{\kappa_{n-d+1}|v_{n-d+1}|^2, \ldots, \kappa_n |v_n|^2\}$ and since the eigenvalues are ordered, the test for a valley point reduces to $V^t \nabla f(x) = 0$ and $\kappa_{n-d+1}(x) < 0$.

This definition disallows the existence of ridges (valleys) of type $n-d$ in regions where W has fewer than d positive (negative) eigenvalues. However, ridges and valleys can occur in hyperbolic regions, where W has both positive and negative eigenvalues. It is possible that a ridge exists at a point where a negative eigenvalue has larger magnitude than any of the positive eigenvalues. A refined definition takes into account the relative magnitudes of the eigenvalues.

- A ridge point x of type $n-d$ is a *strong ridge point* if $\kappa_d(x) > |\kappa_n(x)|$; otherwise, it is a *weak ridge point*. Strong ridge points occur in regions where the convexity dominates the concavity.

- A valley point x of type $n-d$ is a *strong valley point* if $|\kappa_{n-d+1}(x)| > \kappa_1(x)$; otherwise, it is a *weak valley point*. Strong valley points occur in regions where the concavity dominates the convexity.

2.3.4 Examples

Example 2.1: Let $f(x,y) = x^2 y$. The eigenvalues of W are $\kappa_1 = -y + \sqrt{4x^2+y^2}$ and $\kappa_2 = -y - \sqrt{4x^2+y^2}$. Observe that $\kappa_1(x,y) \geq 0 \geq \kappa_2(x,y)$ for all (x,y). Also, $\kappa_1 > |\kappa_2|$ for $y < 0$ and $|\kappa_2| > \kappa_1$ for $y > 0$. The corresponding eigenvectors are

$$v_1 = \left\{ \begin{array}{ll} (-y+R, -2x), & y \leq 0 \\ (2x, -y-R), & y \geq 0 \end{array} \right\}, \quad v_2 = \left\{ \begin{array}{ll} (2x, -y+R), & y \leq 0 \\ (-y-R, -2x), & y \geq 0 \end{array} \right\}$$

where $R = \sqrt{4x^2+y^2}$. The two different sets of eigenvectors are used since at $x = 0$ one vector in each set degenerates to the zero vector. The first directional derivatives are

$$D_{v_1} f = \left\{ \begin{array}{ll} 2x(yR - x^2 - y^2), & y \leq 0 \\ x^2(3y - R), & y \geq 0 \end{array} \right\}, \quad D_{v_2} f = \left\{ \begin{array}{ll} x^2(3y+R), & y \leq 0 \\ -2x(yR + x^2 + y^2), & y \geq 0 \end{array} \right\}.$$

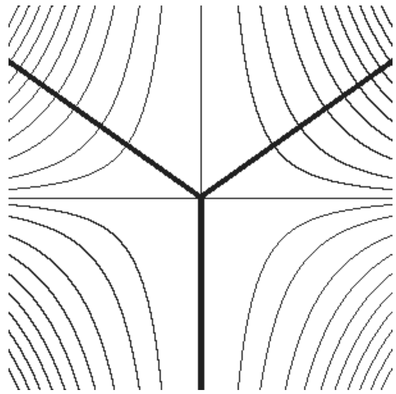


Figure 2.1: Height ridges of x^2y

Let V be the 2×2 matrix whose columns are v_1 and v_2. Since $\mathrm{H}(f)$ is a symmetric matrix, its eigenvectors v_1 and v_2 are orthogonal; thus we have $-V^t\mathrm{H}(f)V = \mathrm{diag}\{\kappa_1|v_1|^2, \kappa_2|v_2|^2\}$.

There are no ridges of type 0 since $\kappa_2 \leq 0$ for all (x, y). However, there are ridges of type 1. Firstly, note that $\kappa_1(x, y) > 0$ as long as not both $x = 0$ and $y \geq 0$. Secondly, if $x = 0$ and $y < 0$, then $D_{v_1}f = 0$ and $D_{v_1v_1}f < 0$. Also, if $y > 0$ and $x^2 = y^2$, then $D_{v_1}f = 0$ and $D_{v_1v_1}f < 0$. Therefore, the ridges of type 1 lie on three rays with origin $(0, 0)$. The ray $x = 0$, $y < 0$ is a strong ridge since $\kappa_1 > |\kappa_2|$ at those points. The rays $x^2 = 2y^2$, $y > 0$ are weak ridges since $\kappa_1 < |\kappa_2|$ at those points.

Figure 2.1 contains a contour plot with the ridges drawn as thick lines. The origin is at the center of the picture and the coordinates are right–handed. □

Example 2.2: Consider $f(x, y, z) = \frac{1}{2}(ax^2 + by^2 + cz^2)$ where $0 < a < b < c$. The first derivatives are $\nabla f = (ax, by, cz)$, and the second derivatives are $\mathrm{H}(f) = \mathrm{diag}(a, b, c)$. The matrix W has ordered eigenvalues $\kappa_1 = -a$, $\kappa_2 = -b$, and $\kappa_3 = -c$, with corresponding eigenvectors $v_1 = (1, 0, 0)$, $v_2 = (0, 1, 0)$, and $v_3 = (0, 0, 1)$. Since the eigenvalues are all negative, valley points are expected, but not ridge points.

The only valley point of type 0 is the local minimum point $(0, 0, 0)$. The valley points of type 1 consist of the x–axis since $D_{v_2}f = D_{v_3}f = 0$ imply $y = z = 0$. Intuitively this seems reasonable since the longest axes of the ellipsoidal level sets lie on the x–axis. The valley points of type 2 consist of the xy–plane since $D_{v_3}f = 0$ implies $z = 0$. This set also makes intuitive sense since the ellipsoidal level sets are flattest in the z–direction. □

2.3.5 Invariance Properties

The ridges constructed by the height definition are invariant under spatial translations, spatial rotations, and uniform spatial magnifications. They are not invariant with respect to monotonic transformations of the intensity function. The following notation is used in the proofs. Let $x \in \mathbb{R}^n$. Let $u = u(x)$ be an invertible change of spatial variables, and let $\bar{f}(u) = f(x)$. Define $du/dx = [\partial u_i / \partial x_j]$ to be the matrix of partial derivatives of the components of u with respect to the components of x. All functions u of the class of spatial transformations mentioned above have the property $\partial^2 u_k / \partial x_i \partial x_j = 0$. Consequently, the following relationships hold:

$$\nabla f = \left(\frac{du}{dx}\right)^t \nabla \bar{f} \;\text{ and }\; \mathrm{H}(f) = \left(\frac{du}{dx}\right)^t \mathrm{H}(\bar{f}) \frac{du}{dx}.$$

Finally, define $W = -\mathrm{H}(f)$ and $\bar{W} = -\mathrm{H}(\bar{f})$. The generic eigenvalues and eigenvectors for these matrices will be denoted by κ, v, $\bar{\kappa}$, and $\bar{v}$, accordingly.

Invariance under spatial translations. Let $u = x + a$ where a is a constant vector; then $\nabla f = \nabla \bar{f}$ and $\mathrm{H}(f) = \mathrm{H}(\bar{f})$. Consequently, $\bar{v} = v$, $\bar{\kappa} = \kappa$, and x_0 is a solution to $D_v f = 0$ if and only if $u_0 = x_0 + a$ is a solution to $D_{\bar{v}} \bar{f} = 0$. The eigenvalue comparison in the strong ridge definition must hold since the eigenvalues have not changed magnitude.

Invariance under spatial rotations. Let $u = Rx$ where R is a rotation matrix; then $\nabla f = R^t \nabla \bar{f}$ and $\mathrm{H}(f) = R^t \mathrm{H}(\bar{f}) R$. Consequently, $\bar{v} = Rv$, $\bar{\kappa} = \kappa$, and x_0 is a solution to $D_v f = 0$ if and only if $u_0 = Rx_0$ is a solution to $D_{\bar{v}} \bar{f} = 0$. The eigenvalue comparison in the strong ridge definition must hold since the rotation does not change the magnitudes of the eigenvalues.

Invariance under uniform spatial magnification. Let $u = cx$ where c is a positive scalar; then $\nabla f = c \nabla \bar{f}$ and $\mathrm{H}(f) = c^2 \mathrm{H}(\bar{f})$. Consequently, $\bar{v} = v$, $\bar{\kappa} = \kappa / c^2$, and x_0 is a solution to $D_v f = 0$ if and only if $u_0 = cx_0$ is a solution to $D_{\bar{v}} \bar{f} = 0$. The eigenvalue comparison in the strong ridge definition must hold since dividing a set of numbers by a positive value does not change the ordering of the set.

Lack of invariance under monotonic transformations. Let $f(x,y) = 1 - x^2 - 2y^2$. The eigenvalues of $-\mathrm{H}(f)$ are $\kappa_1 = 4$ and $\kappa_2 = 2$ with corresponding eigenvectors $v_1 = (0,1)$ and $v_2 = (1,0)$. The point $(1,0)$ is a ridge point since $v_1 \cdot \nabla f(1,0) = 0$ and $\kappa_1(1,0) > 0$. Let $g : \mathbb{R} \to \mathbb{R}$ be a differentiable function such that $g' > 0$. Define the composition of functions $\phi = g \circ f$. The derivatives are related by $\nabla\phi = g'(f)\nabla f$ and $\mathrm{H}(\phi) = g'(f)\mathrm{H}(f) + g''(f)\nabla f \nabla f^t$. The eigenvalues of $-\mathrm{H}(\phi(1,0))$ are $2g'(0) - 4g''(0)$ and $4g'(0)$ with corresponding eigenvectors $(1,0)$ and $(0,1)$, respectively. In order that $(1,0)$ be a ridge point for ϕ it is necessary that $\kappa_1 = 4g'(0) > 2g'(0) - 4g''(0) = \kappa_2$. In this case $v_1 = (0,1)$ and $v_1 \cdot \nabla\phi(1,0) = 0$. But the condition $4g'(0) > 2g'(0) - 4g''(0)$ is not satisfied by all monotonic functions, for example $g(t) = \ln(1+t)$, where $g'(0) = 1$ and $g''(0) = -1$.

For a discussion of second–order invariance properties under general intensity transformations, see Florack, ter Haar Romeny, Koenderink & Viergever (n.d.).

2.4 Principal Direction Definition

The principal direction definition for ridges and valleys is motivated by the differential geometry of n–dimensional hypersurfaces in $\mathbb{R}^{n+1}$. I define creases as loci of extrema of *principal curvatures* along associated *lines of curvature*. The curvature measurements are made with respect to the metric on the tangent hyperplanes. For a geometric motivation, compare with (Scharlach 1993).

In standard differential geometry textbooks, hypersurfaces are described by a parameterization which is used in obtaining principal curvatures and principal directions. For example, if the hypersurface is the graph of a function $f(x)$, it is parameterized by $(x, f(x))$. In some applications there may not be a natural parameterization for the hypersurface of interest. For example, if the skull is segmented in a three dimensional data set as a surface, it is not naturally parameterized by spatial coordinates. But the surface may be thought of as a level surface for a function of three spatial variables. I want to construct principal curvatures and principal directions for surfaces defined as level sets of functions $F : \mathbb{R}^{n+1} \to \mathbb{R}$. Assume that F is a C^4 function for which $\nabla F \neq 0$. The normal vectors to the surface are

$N = \nabla F / |\nabla F|$.

Construction With Parameterization. Let the surface be parameterized by position $x : \mathbb{R}^n \rightarrow \mathbb{R}^{n+1}$, say $x = x(u)$. Define $J = dx/du$, an $(n+1) \times n$ matrix whose entries are the partial derivatives of the components of x with respect to the components of u. The matrix has rank n and satisfies the property $N^t J = 0$. That is, the columns of J are a basis of the tangent space and are orthogonal to N at position $x(u)$. The first and second fundamental forms are given by the $n \times n$ matrices $\mathbf{I} = J^t J$ and $\mathbf{II} = -J^t dN/du$, respectively. The matrix representing the shape operator on the tangent space is $S = \mathbf{I}^{-1}\mathbf{II}$. Consider the eigenvector problem $Sp = \kappa p$. Each eigenvector p is a principal direction. The corresponding eigenvalue κ is a principal curvature. The vector p is an n–vector given in terms of tangent space coordinates, but its representation in $\mathbb{R}^{n+1}$ is $\xi = Jp$.

Construction without Parameterization. Let the surface be defined implicitly by $F(x) = 0$ where $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$. The variable x represents position, but is not parameterized as it was in the previous paragraph. Define $W = -dN/dx$, an $(n+1) \times (n+1)$ matrix whose entries are the partial derivatives of the components of the normal vector N with respect to the components of position x. Note that $N \cdot N = 1$ implies $N^t W = 0$. I claim that if $Sp = \kappa p$, then $\xi = Jp$ satisfies $W\xi = \kappa\xi$. Firstly, note that $\mathbf{I} = J^t J$. Secondly, by the chain rule, $dN/du = (dN/dx)J$, so $\mathbf{II} = J^t W J$. The eigenvector problem $(\mathbf{II} - \kappa\mathbf{I})p = 0$ is therefore transformed to $J^t(W - \kappa E)\xi = 0$, where E is the $(n+1) \times (n+1)$ identity matrix and where $\xi = Jp$.

Since J^t has full rank n, the general solution to $J^t(W - \kappa E)\xi = 0$ is

$$(W - \kappa E)\xi = [E - J(J^t J)^{-1} J^t]c = Mc$$

where $M = E - J(J^t J)^{-1} J^t$ and c is an arbitrary vector (see Pearson (1983) for least square solutions of linear systems using generalized inverses). Observe that $MJ = 0$, so $\mathrm{range}(M) = \mathrm{span}(N)$. Therefore, $\alpha N = Mc = (W - \kappa E)\xi$ for some scalar α. Multiplying on the left by N^t yields

$$\alpha = N^t(W - \kappa E)\xi = -\kappa N^t \xi = 0,$$

where I have used the facts that $N^t W = 0$, $\xi = Jp$, and $N^t J = 0$. Consequently $(W - \kappa E)\xi = 0$. Conversely, it is easy to see that $(W - \kappa E)\xi = 0$ implies $(\mathbf{II} - \kappa\mathbf{I})p = 0$ as long as ξ is a tangent vector, say $\xi = Jp$.

Additionally, $W(x)$ has an eigenvalue which is identically zero for all x, but the corresponding eigenvector is not a tangent vector. This follows from the identity $W = (E - NN^t)\mathrm{H}(F)/|\nabla F|$ which can be derived by explicitly computing $\partial N_i/\partial x_j$ for $N = \nabla F/|\nabla F|$. The eigenvector is $\mathrm{adj}(\mathrm{H}(F))\nabla F$ where adj indicates the adjoint of a matrix, the transpose of the matrix of cofactors of the input matrix. A short computation shows that $W\,\mathrm{adj}(\mathrm{H}(f))\nabla F = 0$.

2.4.1 Creases on Graphs

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^4 function with graph $g : \mathbb{R}^n \to \mathbb{R}^{n+1}$ given by $g(x) = (x, f(x))$. The principal curvatures κ_i and directions p_i, $1 \le i \le n$, are determined by $Sp_i = \kappa_i p_i$ where S is the shape operator described earlier. Assume that the curvatures are ordered as $\kappa_1 \ge \ldots \ge \kappa_n$. In the height definition, I defined creases as relative extrema of a single real–valued function. The principal direction definition is different in that creases will occur as extreme points of each principal curvature with respect to its principal direction. Moreover, the classification of an extreme point will depend on following the integral curves of the principal direction vector field, so the second directional derivative test must be used rather than considering the definiteness of a Hessian matrix. Like the height definition, I will characterize the creases according to the dimension of the manifold we expect when finding roots to equations. I also can refine the definitions to include the concepts of strong and weak creases. In the definition, assume that $1 \le d \le n$.

- The point x is a *ridge point of type* $n-d$ if $\kappa_d(x) > 0$, $D_{p_i}\kappa_i(x) = 0$, and $D_{p_i}D_{p_i}\kappa_i(x) < 0$ for $1 \le i \le d$. Additionally x is a *strong ridge point* if $\kappa_d(x) > |\kappa_n(x)|$; otherwise it is a *weak ridge point*.

- The point x is a *valley point of type* $n-d$ if $\kappa_{n-d+1}(x) < 0$, $D_{p_i}\kappa_i(x) = 0$, and $D_{p_i}D_{p_i}\kappa_i(x) > 0$ for $n-d+1 \le i \le n$. Additionally x is a *strong valley point* if $|\kappa_{n-d+1}(x)| > \kappa_1(x)$; otherwise it is a *weak valley point*.

I briefly contrast the height and principal direction definitions. In the height definition, I searched for the local extrema of a *single* function f whose domain was restricted to a subspace of $\mathbb{R}^n$ (so the search was in *multiple* directions). That is, if V is the $n \times d$ matrix whose columns span the desired subspace and if s is a $d \times 1$ vector– valued parameter, then the search was for extrema of $\phi(s) = f(x + Vs)$ using the standard definition for extrema. The second derivative test involved determining the definiteness of the second derivative matrix for $\phi(s)$ when $s = 0$. In the principal direction definition, the search is for local extrema of *multiple* functions κ_i. Each such function has a *single* direction p_i associated with it, so the construction of extrema is the usual one for functions of a single real variable. That is, if s is a real variable, for each i search for extrema along a path $\xi(s)$ of a function $\phi_i(s) = \kappa_i(\xi(s))$, where the path is determined by $\xi'(s) = p_i(\xi(s))$, $\xi(0) = x$. The second derivative test involves testing the sign of $\phi_i''(0)$.

2.4.2 Creases on Level Surfaces

Let $F : \mathbb{R}^{n+1} \to \mathbb{R}$ be a C^4 function, and consider the hypersurface defined implicitly by $F(x) = 0$. As shown before, a parameterization of the hypersurface is not necessary to construct its principal curvatures $\kappa_i(x)$ and principal directions $\xi_i(x) \in \mathbb{R}^{n+1}$. They are the eigenvalues and eigenvectors of the matrix $W = -dN/dx$, where $N = \nabla F/|\nabla F|$ and $\xi(x)^t N(x) = 0$. In the definition, assume that $1 \le d \le n$. Also, the points $x \in \mathbb{R}^n$ of interest must be solutions to $F(x) = 0$.

- The point x is a *ridge point of type* $n-d$ if $\kappa_d(x) > 0$, $D_{\xi_i}\kappa_i(x) = 0$, and $D_{\xi_i}D_{\xi_i}\kappa_i(x) < 0$ for $1 \le i \le d$. Additionally x is a *strong ridge point* if $\kappa_d(x) > |\kappa_n(x)|$; otherwise it is a *weak ridge point.*

- The point x is a *valley point of type* $n-d$ if $\kappa_{n-d+1}(x) < 0$, $D_{\xi_i}\kappa_i(x) = 0$, and $D_{\xi_i}D_{\xi_i}\kappa_i(x) > 0$ for $n-d+1 \le i \le n$. Additionally x is a *strong valley point* if $|\kappa_{n-d+1}(x)| > \kappa_1(x)$; otherwise it is a *weak valley point.*

2.4.3 Graph Examples

EXAMPLE 2.3: In dimension $n = 1$, the matrix S is 1×1 and its single entry is $\kappa = -f_{xx}/(1 + f_x^2)^{3/2}$. The graph of $f(x)$ is a planar curve whose curvature at $(x, f(x))$ is $\kappa(x)$. Ridges (valleys) are local maxima (minima) of $\kappa(x)$. For example, let $f(x) = x^p$ where p is

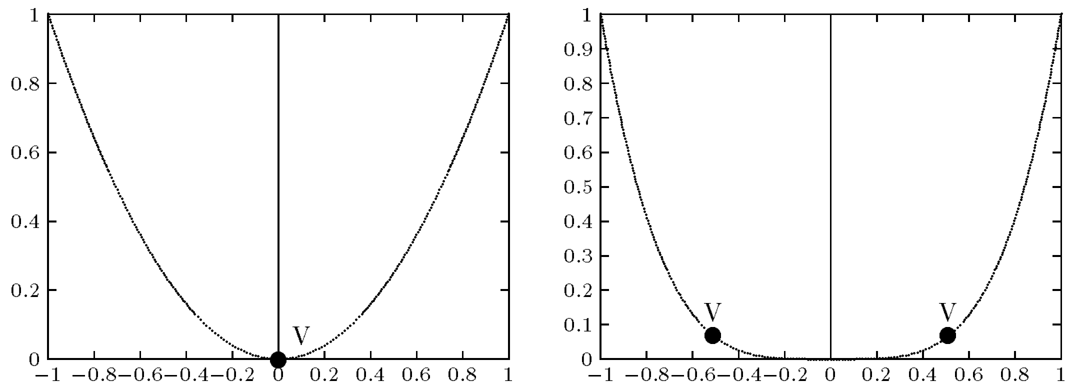


Figure 2.2: Graphs of $f(x) = x^2$ and $f(x) = x^4$

a positive even integer. The curvature is $\kappa = -p(p-1)x^{p-2}/(1+p^2x^{2p-2})^{3/2}$. The solutions to $\kappa_x = 0$ are

$$x = 0,\ \pm\left(\frac{p-2}{p^2(2p-1)}\right)^{1/(2p-2)}.$$

For $p = 2$ the only solution is $x = 0$. The curvature has a negative local minimum of -2, so $x = 0$ is a valley point. For $p > 2$, κ has a local maximum of 0 at $x = 0$, so the graph of f has a flat spot which is neither a ridge nor a valley. At the other two critical points, κ has negative local minima, so the points are valley points. Note that as $p \to \infty$, the graph of x^p approaches 0 pointwise on $(-1, 1)$ and the valley points approach ± 1. Figure 2.2 shows the graphs and valley points for two different values of p. The valley points are labeled on the graphs as V. When $p = 4$, the valleys are $\pm(1/56)^{1/6} \doteq 0.51$. □

The example $f(x) = x^4$ shows that creases according to the principal direction definition are not necessarily local extrema in the function. However, the creases obtained may be better suited for functions which correspond to measurements other than intensity or for which the independent variable is not a spatial one. For example, f might be a function of time for which we are interested in knowing a first time when f has a transition between

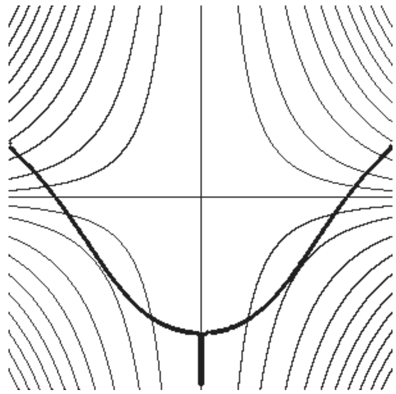


Figure 2.3: Principal direction ridges of x^2y

slowly decreasing and greatly decreasing. The crease points can be viewed as such transitions.

EXAMPLE 2.4: Let $f(x, y) = x^2y$. The matrix for the shape operator is

$$S = \frac{1}{L^3}\begin{bmatrix} -2y(x^4-1) & 2x(1+x^4) \\ 2x(1+2x^2y^2) & -4x^4y \end{bmatrix}$$

where $L = \sqrt{1+|\nabla f|^2} = \sqrt{1+4x^2y^2+x^4}$. The principal curvatures are $\kappa = (y(3x^4-1) \pm R)/L^3$ where $R = \sqrt{(1+x^4)[y^2(9x^4+1)+4x^2]}$, and corresponding principal directions are

$$p_1 = \left\{ \begin{array}{ll} (-(1+x^4)y+R, -2x(1+2x^2y^2)), & y<0 \\ (2x(1+x^4), -(1+x^4)y-R), & y>0 \end{array} \right\}$$

and

$$p_2 = \left\{ \begin{array}{ll} (2x(1+x^4), -(1+x^4)y+R), & y<0 \\ (-(1+x^4)y-R, -2x(1+2x^2y^2)), & y>0 \end{array} \right\}.$$

A closed form solution for the ridges is not tractable. Figure 2.3 shows numerical results for computing the ridges. The figure contains a contour plot with the ridges drawn as thick lines. The origin is at the center of the picture and the coordinates are right–handed. (Compare with the ridges in Figure 2.1). □

2.4.4 Level Surface Example

EXAMPLE 2.5: Consider an ellipsoid defined as a level surface of the function $F(x, y, z) = (ax^2 + by^2 + cz^2)/2$, say $F(x, y, z) = p > 0$, where $0 < a < b < c$. The unit normal vectors are $N = (ax, by, cz)/L$ where $L = \sqrt{a^2x^2 + b^2y^2 + c^2z^2}$. The matrix W is

$$W = \frac{1}{L^3}\begin{bmatrix} a(b^2y^2 + c^2z^2) & -ab^2xy & -ac^2xz \\ -ba^2xy & b(a^2x^2 + c^2z^2) & -bc^2yz \\ -ca^2xz & -cb^2yz & c(a^2x^2 + b^2y^2) \end{bmatrix}.$$

The principal curvatures of the surface are $\kappa_1 = (\alpha + \sqrt{\beta})/L^3$ and $\kappa_2 = (\alpha - \sqrt{\beta})/L^3$ where $\alpha = a^2(b+c)x^2 + b^2(a+c)y^2 + c^2(a+b)z^2$ and $\beta = a^4(b-c)^2x^4 + b^4(a-c)^2y^4 + c^4(a-b)^2z^4 + 2(a-c)(b-c)a^2b^2x^2y^2 + 2(a-b)(c-b)a^2c^2x^2z^2 + 2(b-a)(c-a)b^2c^2y^2z^2$. Corresponding principal directions are

$$p_1 = ab(cxz, cyz, -ax^2 - by^2) + L\kappa_1(acxz, bcyz, -a^2x^2 - b^2y^2)$$

and

$$p_2 = ab(-y(abx^2 + b^2y^2 + c^2z^2), x(a^2x^2 + aby^2 + c^2z^2), c(a-b)xyz) + L^3\kappa_1(-by, ax, 0).$$

Clearly $\kappa_1 > 0$ for all (x, y, z), so I attempt to locate ridges of type 1. Taking derivatives yields the formula

$$L^3\nabla\kappa_1 + 3L^2\kappa\nabla L = \nabla\alpha + \frac{1}{2\sqrt{\beta}}\nabla\beta.$$

At $z = 0$, some calculations will show that $p_1 = -\alpha(0, 0, 1)$. It is easily shown that $p_1 \cdot \nabla\alpha = p_1 \cdot \nabla\beta = p_1 \cdot \nabla L = 0$ when $z = 0$. Thus, $D_{p_1}\kappa_1(x, y, 0) = 0$ for all x and y which lie on the curve $ax^2 + by^2 = 2p$.

The second directional derivative when $z = 0$ can be shown to be $D_{p_1}D_{p_1}\kappa_1 = \alpha[\kappa_{zz} - \alpha\nabla\kappa^t(\partial p_1/\partial z)]$, where all quantities involved are evaluated at $z = 0$. Some tedious algebraic calculations lead to

$$D_{p_1}D_{p_1}\kappa_1(x, y, 0) = \frac{2c^2}{L^3\sqrt{\beta}}\left(\omega_1 a^4x^4 + \omega_2 a^2b^2x^2y^2 + \omega_3 b^4y^4\right)$$

where $\omega_1 = (c-b)[a(4b+5c) - (b+c)(b+6c)]$, $\omega_2 = \{(c-a)[ac + 3(b+c)(b-6c)] + (c-b)[bc + 3(a+c)(a-6c)]\}$ and $\omega_3 = (c-a)[b(4a+5c) - (a+c)(a+6c)]$. Using $0 < a < b < c$, it can be shown that all $\omega_i < 0$, so $D_{p_1} D_{p_1} \kappa_1(x, y, 0) < 0$. Therefore the points on the ellipsoid for which $z = 0$ are ridges of type 1.

More calculations will show that when $y = 0$ and $z = 0$, $D_{p_2}\kappa_2 = 0$ and $D_{p_2} D_{p_2} \kappa_2 < 0$. The vertices $(\pm\sqrt{2p/a}, 0, 0)$ are therefore ridges of type 0. □

2.4.5 Invariance Properties

For parameterized hypersurfaces, the ridge construction is invariant under diffeomorphisms on the parameter space. This result follows from standard differential geometry where the principal curvatures and principal directions do not change under these transformations. With regard to transformations applied to the entire space $\mathbb{R}^{n+1}$ in which the level surface $F(x) = 0$ lives, the ridges constructed are invariant under spatial translations and spatial rotations (Euclidean motions), and under uniform spatial magnifications. The proofs are nearly identical to those in Section 2.3.5.

In the special case of a graph defined by $F(x, z) = z - f(x) = 0$, where $x \in \mathbb{R}^n$ and $z \in \mathbb{R}$, the ridge construction is not invariant to uniform magnification in x. Note that the magnification is not a reparameterization of the original surface; the transformation does change the surface. For example, if $n = 1$ and $f(x) = 1 - x^4$ for $x > 0$, a ridge is $x_0 = 56^{-1/6}$. Let $x = c\bar{x}$ for some $c > 0$ and define $\bar{f}(\bar{x}) = f(x) = 1 - c^4\bar{x}^4$. The ridge for this new function is $\bar{x_0} = (56c^8)^{-1/6} \neq x_0/c$, so the ridge is not invariant.

2.5 Level Definition

2.5.1 Creases on Level Surfaces

Let $f : \mathbb{R}^n \to \mathbb{R}$, $n \geq 2$, be a C^4 function such that $\nabla f \neq 0$ (except at isolated points). The domain of f can be partitioned into its level sets defined by $f(x) = c$ for constants c. Note that a single level set can be viewed as a hypersurface in $\mathbb{R}^n$ implicitly defined by $F(x) = f(x) - c = 0$. Therefore, the principal direction definition may be applied to find creases on the hypersurface for each c in the range of f. The normals for the hypersurface are $N = \nabla F/|\nabla F|$, and the eigenvalues and (tangential) eigenvectors of the matrix $W = -dN/dx$ are the principal curvatures and principal directions. I will construct creases on the graph of f by applying the principal direction definition to each of its level surfaces. The set of all creases of all the level surfaces make up the creases of the graph.

The eigenvalues of W are $\kappa_1 \geq \ldots \geq \kappa_{n-1}$ and 0, with corresponding eigenvectors $\xi_1, \ldots, \xi_{n-1}$ and $\mathrm{adj}(\mathrm{H}(f))\nabla f$. The $\kappa_i(a)$ and $\xi_i(a)$ are the principal curvatures and principal directions for the level surface $f(x) = f(a)$. An attempt can be made to construct crease sets of dimension d where $1 \leq d \leq n-1$. The definition is similar to the principal direction definition, but with one subtle difference. In the principal direction definition, creases were solutions to equations of the type $D_\xi \kappa_i(x) = 0$ where $F(x) = 0$ for a single function F. In the level definition, creases are solutions to the same equations, but now for an entire family of functions $F(x; c) = f(x) - c = 0$.

- The point x is a *ridge point of type* $n-1-d$ if $\kappa_d(x) > 0$, $D_{\xi_i}\kappa_i(x) = 0$, and $D_{\xi_i}D_{\xi_i}\kappa_i(x) < 0$ for $1 \leq i \leq d$. Additionally x is a *strong ridge point* if $\kappa_d(x) > |\kappa_n(x)|$; otherwise it is a *weak ridge point.*

- The point x is a *valley point of type* $n-1-d$ if $\kappa_{n-d}(x) < 0$, $D_{\xi_i}\kappa_i(x) = 0$, and $D_{\xi_i}D_{\xi_i}\kappa_i(x) > 0$ for $n-d \leq i \leq n-1$. Additionally x is a *strong valley point* if $|\kappa_{n-d}(x)| > \kappa_1(x)$; otherwise it is a *weak valley point.*

EXAMPLE 2.6: Consider the case $n = 2$. Normal and tangent vectors to the level curves are given by $N(x,y) = (f_x, f_y)/(f_x^2 + f_y^2)^{1/2}$ and $T(x,y) = (f_y, -f_x)/(f_x^2 + f_y^2)^{1/2}$, and the curvature of the level curves is $\kappa(x,y) = -(f_x^2 f_{yy} - 2f_x f_y f_{xy} + f_y^2 f_{xx})/(f_x^2 + f_y^2)^{3/2}$.

Consider the function $f(x,y) = x^2 y$ for $x > 0$ and $y > 0$. The tangents to level curves are $T = (x, -2y)/(x^2 + 4y^2)^{1/2}$. The curvature and its derivative in the T direction are $\kappa = 6xy/(x^2 + 4y^2)^{3/2}$ and $D_T\kappa = -24xy(x^2 - 5y^2)/(x^2 + 4y^2)^3$. Setting $D_T\kappa = 0$ in the first quadrant yields $x = \sqrt{5}y$, $y > 0$. Some calculations will show that $D_T D_T \kappa < 0$, so the points are ridge points.

Figure 2.4 shows a contour plot for the entire plane. The origin is at the center of the picture, and the coordinates are right–handed. The ridges drawn as thick lines, although the curve $x = 0$ for $y < 0$ is degenerate in that the gradient of f is identically zero on it and $\kappa = 0$. (Compare with the ridges in Figures 2.1 and 2.3). □

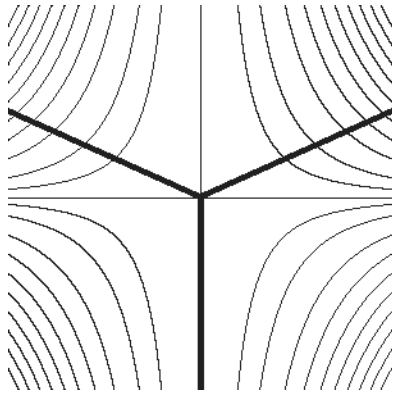


Figure 2.4: Level ridges of $x^2 y$

2.5.2 1–Dimensional Creases from Mean Curvature

In the numerical implementation of the level definition, the eigenvalues and eigenvectors must be computed for the matrix W at each point in an image. This process is typically time–consuming. A variation on the level definition for constructing 1–dimensional creases computes the local extrema of the *mean curvature* $\mu = \mathrm{trace}(W)/(n-1)$ rather than computing local extrema of principal curvatures. The trace of W is more easily computed than its eigenvalues.

Let $f : \mathbb{R}^n \to \mathbb{R}$ and let $x : \mathbb{R}^{n-1} \to \mathbb{R}^n$ denote a level surface; thus, $f(x(s)) \equiv c$ for some constant c and for all s. Let $\phi(s) = \mu(x(s))$ be the mean curvature of the level surface at position $x(s)$. In the construction I use the following abbreviations for the tensor quantities for the gradients and Hessians of the functions of interest:

$$\nabla\phi = \frac{\partial\phi}{\partial s_i}, \nabla\mu = \frac{\partial\mu}{\partial x_k}, \nabla f = \frac{\partial f}{\partial x_k}$$

and

$$\mathrm{H}\phi = \frac{\partial^2\phi}{\partial s_i \partial s_j}, \mathrm{H}\mu = \frac{\partial^2\mu}{\partial x_k \partial x_m}, \mathrm{H}f = \frac{\partial^2 f}{\partial x_k \partial x_m}.$$

I also use the following abbreviations for the tensor quantitites for the first and second derivatives of position $x(s)$:

$$x'(s) = \frac{\partial x_k}{\partial s_i} \quad \text{and} \quad x''(s) = \frac{\partial^2 x_k}{\partial s_i \partial s_j}.$$

The local extrema of ϕ occur when $\nabla\phi = 0$ and $\mathrm{H}\phi$ is positive definite (local maximum) or negative definite (local minimum). I want to determine the local extrema without having to choose a particular parameterization $x(s)$. Select a smoothly varying basis of tangent vectors $v_i(x)$, $1 \le i \le n-1$, which can be used in the derivative tests instead of the tangent vectors $\partial x/\partial s_i$. I will show later that there is such a basis. Let V be the $n \times (n-1)$ matrix whose columns are $v_i(x)$. The two matrices V and $x'(s)$ are related by an invertible $(n-1) \times (n-1)$ matrix C (a change of basis), $x'(s) = VC$.

Using the chain rule, the derivatives of ϕ are

$$\nabla\phi = x'(s)^t \nabla\mu \;\text{ and }\; \mathrm{H}\phi = x'(s)^t\, \mathrm{H}\mu\, x'(s) + x''(s)\nabla\mu.$$

The first derivative test is $0 = \nabla\phi = C^t V^t \nabla\mu$. Since C is invertible, the critical points are solutions to $V^t\nabla\mu = 0$. The second derivative test involves second derivatives of position, which I want to avoid computing. Note that $V^t\nabla\mu = 0$ implies that $\nabla\mu$ is orthogonal to the tangent space of the level surface since the columns of V span the tangent space; that is, $\nabla\mu = \rho\nabla f$ where $\rho = (\nabla\mu \cdot \nabla f)/(\nabla f \cdot \nabla f)$. Moreover, since $f(x(s)) \equiv c$, taking derivatives yields

$$0 = x'(s)^t \nabla f \;\text{ and }\; 0 = x'(s)^t \mathrm{H}f x'(s) + x''(s)\nabla f.$$

Consequently at a critical point,

$$x''(s)\nabla\mu = \rho x''(s)\nabla f = -\rho x'(s)^t \mathrm{H}f x'(s),$$

so the second derivative of ϕ at such points is

$$\mathrm{H}\phi = x'(s)^t\, (\mathrm{H}\mu - \rho\mathrm{H}f)\, x'(s).$$

In terms of the matrix V,

$$C^{-t}\mathrm{H}\phi C^{-1} = V^t\, (\mathrm{H}\mu - \rho\mathrm{H}f)\, V.$$

By Sylvester's Theorem (Horn & Johnson 1991), $C^{-t}\mathrm{H}\phi C^{-1}$ and $\mathrm{H}\phi$ have the same definiteness, so only the definiteness of $V^t(\mathrm{H}\mu - \rho\mathrm{H}f)V$ needs to be checked for the second derivative test.

The remaining problem is to find a smoothly varying basis v_i for the tangent space to level surfaces which is easier to compute than the principal directions. Such a basis is given by the columns of a rotation matrix which maps the vector $e_n = (0, \dots, 0, 1)$ to the normal $N = \nabla f / |\nabla f|$ of the surface. A rotation matrix is given in block form by

$$R = \left[\begin{array}{c|c} E + (N_n - 1)PP^t & Q \\ \hline -Q^t & N_n \end{array}\right]$$

where E is the $(n-1) \times (n-1)$ identity matrix, $Q = (N_1, \dots, N_{n-1})^t$ are the first $n-1$ components of the normal vector, N_n is the last component of N, and $P = Q/|Q|$ when $Q \neq 0$. If $Q = 0$, the rotation matrix is just the identity matrix.

The crease definitions for this variation are given below. Let V be the matrix whose columns are the v_i vectors. A point $x \in \mathbb{R}^n$ is

- a ridge point if $\mu(x) > 0$, $V^t \nabla\mu(x) = 0$, and $V^t(\mathrm{H}\mu - \rho \mathrm{H} f)V$ is negative definite, or
- a valley point if $\mu(x) < 0$, $V^t \nabla\mu(x) = 0$, and $V^t(\mathrm{H}\mu - \rho \mathrm{H} f)V$ is positive definite,

where $\rho(x) = (\nabla\mu \cdot \nabla f)/(\nabla f \cdot \nabla f)$. The directional derivatives and eigenvalues are all evaluated at the point in question. Note that the mean curvature definition is identical to the level definition in the case $n = 2$. The qualitative differences between ridges obtained by the level definition and those obtained by the variation involving mean curvature should be minimal in convex regions (all $\kappa_i > 0$). Some noticeable differences may occur in hyperbolic regions.

2.5.3 Invariance Properties

The ridges constructed by the level definition are invariant with respect to spatial translations, spatial rotations, and uniform spatial magnifications, just as in the principal direction definition since the ridges are located on level surfaces using the principal definition.

The ridges are also invariant under monotonic transformations of the function $f : \mathbb{R}^n \to \mathbb{R}$. Intuitively, think of $\mathbb{R}^n$ as a 1–parameter family of level sets of f, each having its function value as an "attribute". Monotonic transformations on f will not change the geometric structure of the level sets; rather it will only change the attributes of the level sets.

The mathematical proof is straightforward. Let $N[f]$ be the normal vectors corresponding to f and let $W[f] = dN[f]/dx$. Let $g : \mathbb{R} \to \mathbb{R}$ be a smooth increasing function ($g' > 0$), and define $h : \mathbb{R}^n \to \mathbb{R}$ to be the composition $h = g \circ f$. Let $N[h]$ be the normal vectors corresponding to h and let $W[h] = dN[h]/dx$; then

$$N[h] = \frac{\nabla h}{|\nabla h|} = \frac{g'(f(x))\nabla f(x)}{|g'(f(x))\nabla f(x)|} = \frac{\nabla f}{|\nabla f|} = N[f].$$

Consequently $W[h] = W[f]$, and the principal curvatures and principal directions are the same for f and h. The construction of creases is identical in either case.

2.6 Nonmetric Definition

The principal direction definition applied to graphs of functions involved finding those values κ and vectors $p \neq 0$ which solve $\mathbf{II}p = \kappa \mathbf{I}p$ where $\mathbf{I}$ is the first fundamental form and $\mathbf{II}$ is the second fundamental form. The presence of $\mathbf{I}$ means that measurements are made in the tangent spaces to the graph. The nonmetric definition is a variation on the principal direction definition. Consider instead solving $\mathbf{II}v = \kappa v$ where the metric $\mathbf{I}$ of the graph is ignored. This change takes us out of the realm of differential geometry. Only graphs of functions are considered with this definition.

2.6.1 General Dimensions

The crease definitions are essentially those for principal directions. Let $f : \mathbb{R}^n \to \mathbb{R}$. Let κ_i and v_i be the eigenvalues and eigenvectors of $\mathbf{II}$ for $1 \leq i \leq n$. Let the eigenvalues be ordered as $\kappa_1 \geq \cdots \geq \kappa_n$. In the definition, assume that $1 \leq d \leq n$.

- The point x is a *ridge point of type* $n-d$ if $\kappa_d(x) > 0$, $D_{v_i}\kappa_i(x) = 0$, and $D_{v_i}D_{v_i}\kappa_i(x) < 0$ for $1 \leq i \leq d$. Additionally x is a *strong ridge point* if $\kappa_d(x) > |\kappa_n(x)|$; otherwise it is a *weak ridge point.*

- The point x is a *valley point of type* $n-d$ if $\kappa_{n-d+1}(x) < 0$, $D_{v_i}\kappa_i(x) = 0$, and $D_{v_i}D_{v_i}\kappa_i(x) > 0$ for $n-d+1 \leq i \leq n$. Additionally x is a *strong valley point* if $|\kappa_{n-d+1}(x)| > \kappa_1(x)$; otherwise it is a *weak valley point.*

EXAMPLE 2.7: In dimension $n = 1$, the eigenvalue formula is $\kappa = -f_{xx}/(1+f_x^2)^{1/2}$. Consider $f(x) = \exp(-x^2/2)/\sqrt{2\pi}$, a Gaussian distribution with standard deviation 1. The function has a relatively slow decrease for small x. For larger x the graph drops off sharply and then remains close to 0 as $x \to \infty$. The creases obtained from the nonmetric definition will be used as markers of where the rate of decrease of f has transitions from slow to fast.

For this example, $\kappa_x = 0$ when $x_0 = 0$ or when x is a solution to $2\pi(x^2 - 3)\exp(x^2) = x^2 + 1$. Using Newton's method to find an approximate solution yields $x_1 \doteq 1.741$. Also, $\kappa(x_0) = 1$, $\kappa_{xx}(x_0) = -4$, $\kappa(x_1) \doteq -0.176$, and $\kappa_{xx}(x_1) \doteq 0.557$, so x_0 is a ridge point and x_1 is a valley point. □

Example 2.8: In dimension $n = 2$ consider the same function $f(x, y) = x^2 y$ as before, so the ridge definitions can be compared. Define $L = \sqrt{1 + |\nabla f|^2} = \sqrt{1 + 4x^2y^2 + x^4}$ and $R = \sqrt{4x^2 + y^2}$. The eigenvalues of W are $\kappa_1 = (-y + R)/L$ and $\kappa_2 = (-y - R)/L$ with corresponding eigenvectors

$$v_1 = \left\{ \begin{array}{ll} (-y + R, -2x), & y \le 0 \\ (2x, -y - R), & y \ge 0 \end{array} \right\} \text{ and } v_2 = \left\{ \begin{array}{ll} (2x, -y + R), & y \le 0 \\ (-y - R, -2x), & y \ge 0 \end{array} \right\}.$$

The eigenvalues satisfy $\kappa_1(x, y) \ge 0 \ge \kappa_2(x, y)$ for all (x, y). The first directional derivative of κ_1 is

$$D_{v_1}\kappa_1 = \left\{ \begin{array}{ll} 2x(-y + R)g(x, y)/(RL^3), & y \le 0 \\ 4x^2 g(x, y)/(RL^3), & y \ge 0 \end{array} \right\}$$

where $g(x, y) = 3 - (x^2 - y^2)(x^2 - 2y^2) + Ry(5x^2 + 2y^2)$.

A first set of solutions to $D_{v_1}\kappa_1 = 0$ is given by the line $x = 0$. A second set of solutions can be constructed by setting $y = mx$ in the equation $g(x, y) = 0$. The curve of solutions for $x > 0$ is given parametrically by

$$x(m) = \left[\frac{3}{(1 - m^2)(1 - 2m^2) - m(5 + 2m^2)\sqrt{4 + m^2}} \right]^{1/4}, \quad y(m) = mx(m),$$

where $m \le m_0$ and m_0 is the positive solution to $52m^6 + 92m^4 + 106m^2 - 1 = 0$. The solutions for $x < 0$ are obtained by reflection through the y–axis. Note that $\lim_{m\to-\infty} x(m) = 0$ and $\lim_{m\to-\infty} y(m) = -(3/4)^{1/4}$.

It can be shown by the second derivative test that the points (x, y) for which $x = 0$ and $y < -(3/4)^{1/4}$ are ridge points, and all points $(x(m), y(m))$ given above are ridge points.

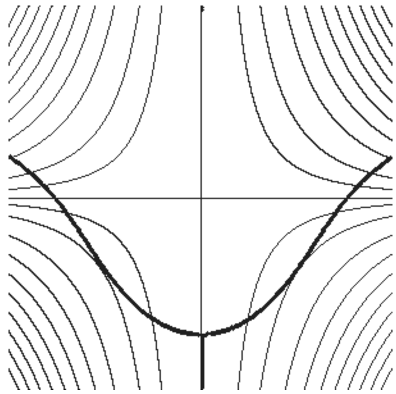


Figure 2.5: Nonmetric ridges of x^2y

Figure 2.5 shows a contour plot with the ridges drawn as thick lines. The origin is at the center of the picture and the coordinates are right–handed. (Compare with the ridges in Figures 2.1, 2.3, and 2.4). □

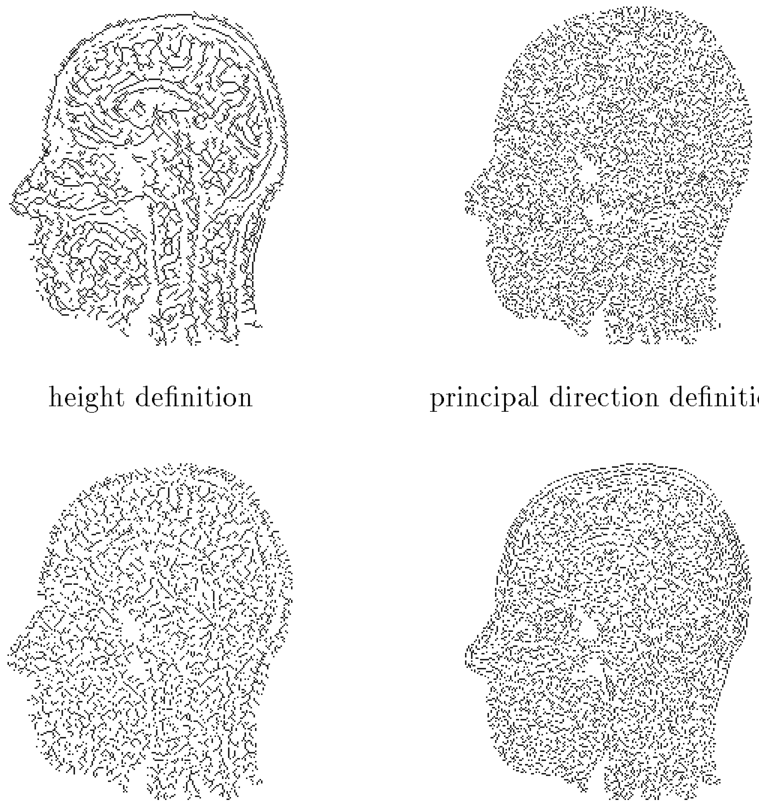
2.6.2 Invariance Properties

The ridges constructed by the nonmetric definition have the same invariance as those of the principal direction definition; the proofs are similar. The ridges are invariant under spatial translations and spatial rotations (in $\mathbb{R}^n$). The ridges are not invariant under uniform spatial magnifications and monotonic transformations of f.

2.7 Experiments on Images

I tested the four basic ridge definitions (height, principal direction, level, and nonmetric) on a slightly blurred version of the MR image shown in Figure 1.4. The original image was assumed to have a (inner) scale value $\sigma = 1$. The image was Gaussian blurred to scale $\sigma = 2$ to help remove small scale noise. I ignored the valleys for clarity in the resulting images in Figure 2.6.

The ridges were constructed by finding zero–crossings of the appropriate directional derivatives. Both pixels involved in a zero–crossing were marked as ridges. The resulting



height definition | principal direction definition
level definition | nonmetric definition

Figure 2.6: Ridges of MR head image

binary set was thinned using morphological operations. The ridges from the height definition appear to give the best qualitative information about structure of the image. The ridges from both the principal direction and nonmetric definitions seem to be overly abundant. The level definition appears to produce more dendritic–like ridge structures, but the ridges are fragmented. The reduction in detail through increasing scale seems to happen sooner for the height definition than for the other definitions. In the principal direction case, no object structures are apparent at the selected scale. In the nonmetric case you can see the scalp, corpus colosum, and brain stem.

2.8 Discussion

The four crease definitions: height, principal direction, level, and nonmetric, appear to produce qualitatively different structures. When applied to graphs of functions all four provide invariance under rotations and translations in the spatial variables. The level definition is additionally invariant under monotonic transformations of the function values. Based on the images of Figure 2.6, the height definition seems to capture the most large scale shape information about the image. The dendritic structures occurring from the level definition seem to be suitable for providing small scale descriptions of objects in the image.

The principal direction and nonmetric definitions applied to graphs do not provide invariance under uniform spatial magnifications and monotonic changes in intensity. Moreover, the fact that the ridges tend not to correspond to object centers makes these definitions not suitable for medical image analsis. However the principal direction definition does apply to surfaces whose coordinates all correspond to spatial information, so the definition is suitable for locating ridges on surfaces and using them for 3D image registration.

The true test of the usefulness of ridges in image analysis lies in multiscale analysis. Ridge structure at a single scale can give information about structures of an appropriate size, but the behavior of ridge structures as the scale parameter is increased allows one to capture more global object information. An application of ridge analysis to image segmentation using a multiscale approach is given in Chapter 3. The basis for applying ridge analysis to medial analysis based on cores is given in Chapter 4.