THE INTRINSIC GEOMETRY OF DYNAMIC OBSERVATIONS

Arthur J. Krener

ABSTRACT

There are several ways to introduce geometry into the problem of estimating the state of nonlinear process given observations of it. We classify these as intrinsic or extrinsic. We show how the linearizability of this problem is related to the existence of an intrinsic Koszul connection on the output space and its curvature and torsion.

1. Extrinsic Geometry

Consider the problem of estimating a process $\xi(t)$ from observation of a related process $\psi(t)$. This can be formulated in stochastic terms as a nonlinear filtering problem. We assume that the two processes are described by stochastic differential equations

$$d\xi = f(\xi)dt + g(\xi)dw \qquad (1.1a)$$

$$d\psi = h(\xi)dt + k(\xi)dv \qquad (1.1b)$$

$$\xi(t_0) = \xi^0 \qquad (1.1c)$$

The state process $\xi(t)$ and output process $\psi(t)$ evolve on n and p dimensional manifolds N and P. The driving processes $w(t)$ and $v(t)$ are m and p dimensional independent standard Wiener processes. The initial condition ξ^0 is an N valued random variable independent of $w(t)$ and $v(t)$. This (1.1) is a local coordinate description using Ito differentials. We regularly abuse notation by confusing local coordinate descriptions with the intrinsic objects they describe.

The nonlinear filtering problem is to compute in real time the conditional distribution of the current state $\xi(t)$ (or some useful statistics such as the conditional mean) given the past observations $\psi(s)$, $t_0 \leq s \leq t$. We assume that f, h, g, k and the distribution of ξ^0 are known. This is an extremely important and extremely difficult problem. Kalman and Bucy discovered the only broad class of models for which an efficient algorithm is known. These are the linear filtering problems.

$$dx = A(t)x\,dt + B(t)dw \qquad (1.2a)$$

$$dy = C(t)x\,dt + D(t)dv \qquad (1.2b)$$

$$x(t_0) = x^0 \qquad (1.2c)$$

Research supported in part by the NSF under MSC-8300884

M. Fliess and M. Hazewinkel (eds.), Algebraic and Geometric Methods in Nonlinear Control Theory, 77–87.

These can be viewed as a special case of (1.1) where $\xi = (t,x)$ and $\psi = y$.

If the initial condition x^0 is Gaussian then the conditional distribution of $x(t)$ is also Gaussian hence completely described by its mean and covariance. These evolve according to ordinary differential equations and so are easily computable in real time. Actually only the conditional mean need be computed in real time for the conditional variance evolves independently of the observations

There have been several attempts to relate the complexity of the nonlinear filtering problem to certain geometric aspects of the model (1.1). (We use the term geometry in the broadest sense to include Lie theoretic concepts.)

The program which has received the greatest effort was initiated by Brockett [1] and recently surveyed by Marcus [3]. The basic goal is to find models (1.1) other than (1.2) for which the evolution of the conditional density of $\xi(t)$ or some useful statistic such as the conditional mean could be described by a finite number of sufficient statistics which evolve according to ordinary differential equations driven by ψ.

The approach is to try to reduce the partial differential equation which describes the evolution of the conditional density to ordinary differential equations for a set of sufficient statistics. A certain Lie algebra can be associated to the model (1.1) and if this algebra is finite dimensional or admits an ideal of finite codimension then the program can be successfully carried out. Unfortunately this approach has not led to any new broad class of finite dimensional filters.

With the benefit of hindsight there are several criticisms that one can make of this program. We should not really expect the conditional density of $\xi(t)$ to be any easier to describe than the unconditional density of $\xi(t)$ for the former can be seen as a special case of the latter when $h(\xi)$ is constant. Yet (1.2) is the only broad class of models where the evolution of the unconditional density from a family of initial distributions can be described by a finite number of sufficient statistics. Moreover virtually nothing is known regarding infinite dimensional nonlinear realization theory. The finite dimensional theory has been extensively studied. This depends heavily on invariant foliations rather than finite dimensional Lie algebras. Therefore it seems a bit naive to expect infinite dimensional Lie algebras to play such an important role in the infinite dimensional case.

Finally the Lie algebraic structure is extremely sensitive to small perturbations of the problem. A small change in f, g, h and k of (1.1) can dramatically change the resulting algebra. For example, a small cubic nonlinearity can transform the Lie algebra of (1.2) from low dimensional to infinite dimensional [3].

The last point is particularly undesirable because the essential role g and k play in defining the Lie algebra. Typically f and h are derived from physical laws. The directions of g and k may also follows from such laws but the magnitude and their dependence on ξ are usually no more than the result of educated guesses by the modeler. Even for the linear case (1.2) there is no consensus on how to choose B and D.

It is for this reason that we call extrinsic any geometric structure associated to (1.1) which depends on how the observations are imbedded in noise, i.e., on g and k. If the structure only depends on the observations, i.e., only on f and h, then we call it intrinsic.

Another example of an extrinsic geometric structure is as follows. We can define a vector bundle E over N by letting the fiber E_ξ be the vector space.

$$E_\xi = \text{Col Span} \begin{bmatrix} g(\xi) & 0 \\ 0 & k(\xi) \end{bmatrix} \tag{1.3}$$

We assume that this matrix is everywhere of full rank so that the fiber dimension is m+p. Since the Wiener processes w(t) and v(t) are standard it is natural to think of the columns of (1.3) as defining an orthonormal basis for E_ξ, and hence a Riemannian structure on E. This construction is even less intrinsic than the previous because we have used only $g(\xi)$ and $k(\xi)$.

There have been suggestions in the literature that a study of this Riemannian structure could assist in understanding the nonlinear filtering problem. Perhaps this might be so but the arbitrariness in the modeler's choice of g and k makes this approach highly suspect.

2. Intrinsic Geometry

We will construct from f and h something resembling a Koszul connection on the output space and show how the existence of this connection, its torsion and curvature relate to the difficulty of the observation problem. Recall [4] that a Koszul connection ∇ is mapping from pairs of vector fields to a third vector field.

$$\nabla : (X,Y) \mapsto \nabla_X Y. \tag{2.1}$$

This mapping is linear over C^∞ functions in the first argument X and satisfies a Liebnitz formula in the second argument Y. If $a_i(\psi)$ and $b_j(\psi)$ are C^∞ functions then

$$\nabla_{\sum_i a_i X^i} (\sum_j b_j Y^j) = \sum_{i,j} (a_i b_j \nabla_{X^i} Y^j + a_i L_{X^i}(b_j) Y^j). \tag{2.2}$$

$L_{X^i}(b_j)$ denotes Lie differentiation of b_j by X^i. From this it follows that ∇ is completely determined locally by its Christoffel symbols Γ_k^{ij} where

$$\nabla_{\frac{\partial}{\partial \psi_i}} \frac{\partial}{\partial \psi_j} = \Sigma \, \Gamma_k^{ij} \frac{\partial}{\partial \psi_k} . \tag{2.3}$$

The simplest example of a Koszul connection is $\mathbb{R}^p$ with $\Gamma_k^{ij} = 0$. This is called the linear connection on $\mathbb{R}^p$ because it depends on the linear structure of $\mathbb{R}^p$. Intuitively $\nabla_X Y$ can be thought how Y is twisting along integral curves of X. If $\nabla_X Y = 0$ then Y is parallel along the integral curves of X. In this way ∇ defines parallel translation on P and "connects" the various tangent spaces.

In defining the Christoffel symbols (2.3) we could have taken any frame of vector fields not necessarily a coordinate frame. For our purposes the latter will suffice. Of course (2.3) is a coordinate dependent description of ∇. If we make a change of coordinates $\tilde{\psi} = \tilde{\psi}(\psi)$ then the new Christoffel symbols are given by

$$\tilde{\Gamma}_k^{ij} = \sum_{\rho,\sigma,\tau} \Gamma_\tau^{\rho\sigma} \frac{\partial\psi_\rho}{\partial\tilde{\psi}_i} \frac{\partial\psi_\sigma}{\partial\tilde{\psi}_j} \frac{\partial\tilde{\psi}_k}{\partial\psi_\tau} + \sum_\tau \frac{\partial^2\psi_\tau}{\partial\tilde{\psi}_i\,\partial\tilde{\psi}_j} \frac{\partial\tilde{\psi}_k}{\partial\psi_\tau} \tag{2.4}$$

Because of the second summation, Γ_k^{ij} cannot be the components of a tensor. Moreover it is possible for the Christoffel symbols to be zero with respect to one coordinate system but not another.

We return to our dynamic observation problem. Since we are ignoring the coefficients of the noise, we write it in a deterministic form

$$\dot{\xi} = f(\xi) \tag{2.5a}$$

$$\psi = h(\xi) \tag{2.5b}$$

To be completely analogous to (1.1) we should put a dot over the ψ in (2.5b), but this notation is more standard and is essentially equivalent.

We assume that (2.5) is <u>observable</u> with all <u>observability indices</u> equal to $\ell = n/p$. (This greatly simplifies the analysis. The general case is treated in [2] from a different point of view. We discuss this later.)

The observability assumption means that the functions $L_f^{k-1}(\psi_i)$, for $i = 1, \dots, p; k = 1, \dots, \ell$, are independent and hence could be taken as local coordinates on N.

We define p vector fields g^j, $j = 1, \dots, p$ by

$$L_{g^j} L_f^{k-1}(\psi_i) = \begin{cases} 0 & \text{for } 1 \le k < \ell \\ \delta_i^j & \text{for } k = \ell \end{cases} \tag{2.6}$$

From this we define Γ_k^{ij} by

$$\Gamma_k^{ij} = \frac{1}{\ell} \langle L_f(d\psi_k), [\mathrm{ad}^{\ell-1}(-f)g^i, \mathrm{ad}^{\ell-2}(-f)g^j] \rangle \tag{2.7}$$

Generally these are functions on N but they transform like Christoffel symbols under change of coordinates on P.

<u>Lemma 1</u>. Let $\tilde{\psi} = \tilde{\psi}(\psi)$ be a change of coordinates on P and let $\tilde{g}^j$, $\tilde{\Gamma}_k^{ij}$ be defined by (2.6) and (2.7) in the $\tilde{\psi}$ coordinates.

Proof. By induction for $k = 1,\dots,\ell$

$$\text{Row Span}\begin{bmatrix} d\psi \\ \vdots \\ L_f^{k-1}(d\psi)\end{bmatrix} = \text{Row Span}\begin{bmatrix} d\tilde{\psi} \\ \vdots \\ L_f^{k-1}(d\tilde{\psi})\end{bmatrix} \tag{2.8}$$

and

$$L_f^k(d\psi) = \frac{\partial\psi}{\partial\tilde{\psi}} L_f^k(d\tilde{\psi}) \quad \text{Mod } (2.8) \tag{2.9}$$

therefore

$$\tilde{g} = g\frac{\partial\psi}{\partial\tilde{\psi}} \tag{2.10}$$

and

$$ad^{\ell-1}(-f)\tilde{g} = ad^{\ell-1}(-f)g\frac{\partial\psi}{\partial\tilde{\psi}} - (\ell-1)\, ad^{\ell-2}(-f)g \; L_f\left(\frac{\partial\psi}{\partial\tilde{\psi}}\right) \tag{2.11}$$

$$\text{Mod } \{g,\dots, ad^{\ell-3}(-f)g\}$$

Now

$$\tilde{\Gamma}_k^{ij} = \frac{1}{\ell} < L_f(d\tilde{\psi}_k), [ad^{\ell-1}(-f)\tilde{g}^i, ad^{\ell-2}(-f)\tilde{g}^j]>$$

We expand this using (2.11) and (2.9). Most of the terms are zero by (2.6). We are left with

$$\tilde{\Gamma}_k^{ij} = \frac{1}{\ell}\sum_{\rho,\sigma,\tau} < \frac{\partial\tilde{\psi}_k}{\partial\psi_\tau} L_f(d\psi_\tau), \tag{2.12}$$

$$[ad^{\ell-1}(-f)g^\rho, ad^{\ell-2}(-f)g^\sigma]\frac{\partial\psi_\rho}{\partial\tilde{\psi}_i}\frac{\partial\psi_\sigma}{\partial\tilde{\psi}_j} +$$

$$ad^{\ell-2}(-f)g^\sigma L_{ad^{\ell-1}(-f)\tilde{g}^i}\left(\frac{\partial\psi_\sigma}{\partial\tilde{\psi}_j}\right)\frac{\partial\psi_\rho}{\partial\tilde{\psi}_i} +$$

$$(\ell-1)\, ad^{\ell-2}(-f)\tilde{g}^\sigma L_{ad^{\ell-2}(-f)\tilde{g}^j} L_f\left(\frac{\partial\psi_\rho}{\partial\tilde{\psi}_i}\right) L_f\left(\frac{\partial\psi_\rho}{\partial\tilde{\psi}_i}\right) >$$

If $\varphi = \varphi(\tilde{\psi})$ then (2.6) implies that

$$L_{ad^{\ell-1}(-f)\tilde{g}^i}(\varphi) = \frac{\partial\varphi}{\partial\tilde{\psi}_i}$$

$$L_{ad^{\ell-2}(-f)\tilde{g}^j}(L_f(\varphi)) = L_{ad^{\ell-1}(-f)\tilde{g}^j}(\varphi) = \frac{\partial \varphi}{\partial \tilde{\psi}_j}$$

Therefore (2.12) reduces to (2.4) as desired.

If (2.7) are only functions of ψ, then they are the Christoffel symbols of a Koszul connection on P intrinsically defined by the dynamic observations (2.5).

Example. Consider the linear dynamic observations

$$\dot{x} = Ax \tag{2.14a}$$
$$y = Cx \tag{2.14b}$$

This is the simplest problem for if (C,A) is an observable pair it is easy to construct an asymptotic observer of x(t).

$$\dot{\hat{x}} = (A + GC)\hat{x} - Gy \tag{2.15a}$$

with error ($\tilde{x} = x-\hat{x}$) dynamics

$$\dot{\tilde{x}} = (A + GC)\tilde{x}. \tag{2.15b}$$

The observability assumption implies that the spectrum of (A+GC) can be arbitrarily determined (up to invariance under complex conjugation) by proper choice of G. Therefore, the error can be made to decay exponentially fast at an arbitrary rate.

We use B^j to denote the vector fields defined by (2.6) applied to (2.14) which in this case reduce to

$$C_i A^{k-1} B^j = \begin{cases} 0 & 1 \leq k < \ell \\ \delta_i^j & k = \ell \end{cases} \tag{2.16}$$

Then

$$\Gamma_k^{ij} = \frac{1}{\ell} < C_k A, [A^{\ell-1}B^i, A^{\ell-2}B^j] > 0$$

so the linear dynamics observations (2.14) induces the linear connection on the output space $P = \mathbb{R}^P$.

We can generalize this example by adding output injection to the dynamics

$$\dot{x} = Ax + J(y) \tag{2.18a}$$

$$y = Cx \tag{2.18b}$$

We can add the same output injection to the observer to obtain the same error dynamics

$$\dot{\hat{x}} = (A + GC)\,\hat{x} - Gy + J(y) \tag{2.19a}$$

$$\dot{\tilde{x}} = (A + GC)\,\tilde{x} \tag{2.19b}$$

It is a straightforward exercise to verify that the vector fields g^j defined by (2.6) applied to (2.18) are the B^j defined by (2.16). In other words, the linear output injected dynamics observations (2.18) induces the same linear connection on $P = \mathbb{R}^p$ as does the linear dynamic observations (2.14).

Suppose we ask the question of when the nonlinear dynamic observations (2.5) and the linear output injected dynamic observations (2.18) are equivalent under changes of coordinates $\xi = \xi(x)$ and $\psi = \psi(y)$. A change of state coordinates leaves Γ^{ij}_k invariant. Under a change of output coordinates the Γ^{ij}_k transform like Christoffel symbols. Clearly a necessary condition for the two problems to be equivalent is that the Γ^{ij}_k of (2.5) be functions of ψ alone, and hence define a connection on P. Furthermore there must be a change of output coordinates $\tilde{\psi} = \tilde{\psi}(\psi)$ which takes the Γ^{ij}_k to $\tilde{\Gamma}^{ij}_k = 0$.

When such a change of coordinates exists is a question at the very heart of geometry. The solution dates to Riemann's Habilitation Lecture of 1854 and the paper he submitted in 1861 to the Paris Academy [4]. Reimann was concerned with the question of when what we now call a Riemannian metric could be transformed to the standard Euclidean metric by change of coordinates. The solutions to the two problems are essentially the same.

We denote the Jacobian of the change of output coordinates by

$$\Phi = (\varphi^j_i)$$

$$\frac{\partial \psi_i}{\partial \tilde{\psi}_j} = \varphi^j_i \tag{2.20}$$

This should be an n x n invertible matrix. Plugging into this (2.4) with $\tilde{\Gamma}^{ij}_k = 0$ we obtain a system of linear first order partial differential equations for the φ^j_i thought of as functions of ψ,

$$\frac{\partial}{\partial \psi_i} \varphi^m_k = -\Sigma\, \Gamma^{ij}_k\, \varphi^m_j$$

The equations (2.20) and (2.21) are first order linear partial differential equations for the desired change of coordinates. Moreover we can address the solvability of (2.21) independently of (2.20). The former are solvable if certain integrability conditions are satisfied (the mixed partials must commute). These are given by

$$\sum_{m,r} \left(\Gamma^{ir}_{k}\,\Gamma^{jm}_{r} - \Gamma^{jr}_{k}\,\Gamma^{im}_{r} + \frac{\partial}{\partial \psi_i}\Gamma^{jm}_{k} - \frac{\partial}{\partial \psi_j}\Gamma^{im}_{k}\right)\varphi^{s}_{m} = 0 \qquad (2.22)$$

Since φ^{s}_{m} is assumed to be invertible, this is equivalent to $R^{ijm}_{k} = 0$ where R^{ijm}_{k} is the coefficient of φ^{s}_{m} in (2.22). The R^{ijm}_{k} are components of the curvature tensor associated to ∇. If they are zero then ∇ is said to be flat.

If ∇ is flat so that (2.21) is solvable then (2.20) is solvable iff the columns of Φ are commuting vector fields. After a little calculation this is seen to be equivalent to

$$T^{ij}_{k} := \Gamma^{ij}_{k} - \Gamma^{ji}_{k} = 0 \qquad (2.23)$$

The T^{ij}_{k} so defined are the components of the torsion tensor associated to ∇. If they are zero then ∇ is said to be torsion free.

Therefore a necessary condition for the nonlinear (2.5) and linear, output injected (2.18) dynamic observations to be equivalent is that (2.7) define the Christoffel symbols of a flat and torsion free Koszul connection on P. However this is not sufficient, we need some additional conditions. Suppose we have transformed output coordinates to $\tilde{\psi} = \tilde{\psi}(\psi)$, so that the Christoffel symbols are zero; $\Gamma^{ij}_{k} = 0$. If this can be transformed into (2.18) where the Christoffel symbols are also zero then from (2.4) we see that $\tilde{\psi}$ is necessarily an affine function of y; $\partial\tilde{\psi}/\partial y$ = constant. Applying the argument of Lemma 1 we see that

$$\tilde{g} = \frac{\partial \xi}{\partial x}\, B\, \frac{\partial y}{\partial \tilde{\psi}}$$

and so

$$\{\mathrm{ad}^{\ell-k}(-f)\,\tilde{g}^{j} : 1 \le k \le \ell;\ 1 \le j \le p\} \qquad (2.24)$$

must be a commuting frame. We call (2.24) the frame of basic vector fields associated to the output map $\tilde{\psi}$.

On the other hand if the basic vector fields (2.24) are a commuting frame then we can choose state coordinates x so that they are a coordinate vector fields. It is straightforward to verify in these state coordinates x and output coordinates $y = \tilde{\psi}$ the nonlinear dynamic observations (2.5) are transformed to (2.18). Therefore we have proved a result of Krener and Respondek [2] which we can restate as follows.

Theorem 2. Let the dynamic observations (2.5) be observable with one distinct observability index ℓ of multiplicity p. It can be transformed into linear, output injected dynamic observations (2.18) iff

(i) The Γ_k^{ij} of (2.7) are functions of ψ.

(ii) The Koszul connection ∇ on P defined by Γ_k^{ij} is flat and torsion free

(iii) The basic vector fields (2.24) corresponding to any output coordinates where the Christoffel symbols are zero must be a commuting frame.

Remark. If an addition $\{ad^{\ell}(-f)\tilde{g}^j : j = 1,\dots,p\}$ are commuting and they commute with the basic vector fields then the output injection $J(y) = 0$ so the system (2.5) can be transformed to a linear on (2.14).

The dynamic observations (2.5) are observable with observability indices $\ell,\dots,\ell_p$ if

(i) $\ell_1 \geq \ell_2 \dots \geq \ell_p$ and $\ell_1+\dots+\ell_p = n$

(ii) After reordering the output coordinates the n function $\{L_f^{k-1}(\psi_1) : 1 \leq i \leq p;\ 1 \leq k \leq \ell_i\}$ are independent hence coordinates.

(iii) If $(k_1,\dots k)$ also satisfy (i) and (ii) then $(\ell_1,\dots\ell_p)$ is less than or equal to $(k_1,\dots,k_p)$ in the lexographic order.

The generalization of the foregoing to dynamic observations with several observabilty indices is not at all straightforward. The main difficulty is proving the analog of Lemma 1. which allows the definition of something like Koszul connection on P. To a certain extent these difficulties can be sidestepped for those dynamic observations (2.5) which are equivalent to linear, output injected observations. The following paraphases [2].

As before we define p vector fields $g^1,\dots,g^p$ by

$$L_{g^j}\left(L_f^{k-1}(\psi_i)\right) = \begin{cases} 0 & 1 \leq k < \ell_i \\ \delta_i^j & k = \ell_i \end{cases}$$

The output coordinates are said to be special if $\ell_j > \ell_i$ implies that

$$L_{g^j}\left(L_f^{k-1}(\psi_i)\right) = 0 \quad \text{for } 1 \leq k \leq \ell_j. \qquad (2.26)$$

Not every nonlinear dynamic observations (2.25) admits special output coordinates. The linear part (2.14) of the linear, output injected dynamic observations (2.18) can always be brought to dual

Brunovsky form by linear change of coordinates and linear output injection. In this form the output coordinates are special. Hence any system transformable to (2.18) admits special output coordinates.

If $\tilde{\psi}$ are special output coordinates for (2.5) then they must satisfy

$$\langle d\tilde{\psi}_i,\ ad^{\ell_j-1}(-f)g^j\rangle = 0 \qquad (2.27)$$

for $\ell_j > \ell_i$. This is an underdetermined system of PDE's for $\tilde{\psi}$. To be solvable first of all the p dimensional column vectors

$$Y^j = \psi_*\left(ad^{\ell_j-1}(-f)g^j\right) \qquad (2.28)$$

should be functions of ψ not ξ and hence define vector fields on the output space. Then by the Frobenius Theorem $\tilde{\psi}$ satisfying (2.27) exists iff the distributions

$$C^\infty\text{span } \{Y^j:\ \ell_j > \ell_i\}$$

are involutive for i = 1,...p.

The important point about special output coordinates is that transformations between such coordinates are necessarily block upper triangular.

In other words, if both ψ and $\tilde{\psi}$ are special output coordinates then

$$\frac{\partial\psi_i}{\partial\tilde{\psi}_j} = 0 \qquad \text{if } \ell_i < \ell_j$$

Theorem 4. (Krener and Respondek [2]) Let the dynamic observations (2.5) be observable with indices $\ell_1 \geq \ldots \geq \ell_p$. It can be transformed into linear, output injected dynamic observations (2.18) iff

(o) Formula (2.28A) defines vector fields on P and the distributions (2.28b) are involutive. Hence special output coordinates exist.

(i) If ψ are special output coordinates then

$$\Gamma_k^{ij} = \frac{1}{\ell_k} < L_f(d\psi_k),\ [ad^{\ell_i-1}(-f)g^i,\ ad^{\ell_j-2}(-f)g^j] >$$

are functions of ψ, hence the Christoffel symbols of a connection on P.

(ii) This connection is flat and torsion free. Hence Γ_k^{ij} can be transformed to zero by change of output coordinates.

(iii) If ψ are special output coordinates in which the $\Gamma_k^{ij} = 0$ then the frame of basic vector fields

$$\{ad^{\ell_j - k}_{(-f)} g^j : \ 1 \le j \le p; \ 1 \le k \le \ell_j\}$$

is commuting.

References

[1] Brockett, R.W. Remarks on finite dimensional nonlinear estimation. Asterique, 75-76 (1980) pp 47-55.

[2] Krener, A.J. and W. Respondek, Nonlinear observers with linearizable error dynamics, to appear, SIAM J. Control and Optimization, 1985.

[3] Marcus, S.I., Algebraic and geometric methods in nonlinear filtering, SIAM J. Control and Optimization 22 (1984) pp 817-844.

[4] Spivak, M. A Comprehensive Introduction to Differential Geometry, V. II, Publish or Perish Press,Berkeley, 1979.