Nonlinear Signal Processing (2004/2005)

jxavier@isr.ist.utl.pt

$\boxed{\textbf{Submanifolds}}$

Theorem [Inverse Function Theorem] Let U and V be open subsets of $\mathbb{R}^n$ and $F : U \to V$ a smooth map. Let $p \in U$. If
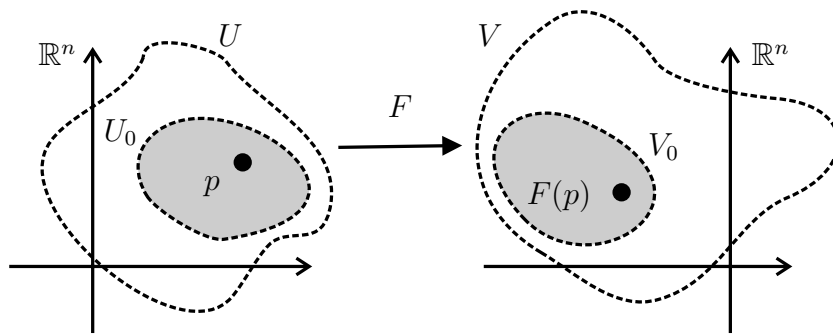
$$
DF(p) = \begin{bmatrix}
\dfrac{\partial F^1}{\partial x^1}(p) & \dfrac{\partial F^1}{\partial x^2}(p) & \cdots & \dfrac{\partial F^1}{\partial x^n}(p) \\[2ex]
\dfrac{\partial F^2}{\partial x^1}(p) & \dfrac{\partial F^2}{\partial x^2}(p) & \cdots & \dfrac{\partial F^2}{\partial x^n}(p) \\[2ex]
\vdots & \vdots & \cdots & \vdots \\[2ex]
\dfrac{\partial F^n}{\partial x^1}(p) & \dfrac{\partial F^n}{\partial x^2}(p) & \cdots & \dfrac{\partial F^n}{\partial x^n}(p)
\end{bmatrix}
$$

is nonsingular, then there exist neighborhoods $U_0 \subset U$ of p and $V_0 \subset V$ of $q = F(p)$ such that $F : U_0 \to V_0$ is a diffeomorphism. Furthermore, we have

$$
DF^{-1}(y_0) = \left(DF\left(x_0 \right) \right)^{-1},
$$

where $x_0 = F^{-1}(y_0)$, for each $y_0 \in V_0$.

$\triangleright$ *Intuition: the bijectivity of $DF(p)$ carries over $\underline{locally}$ to F*

Example 1 [Inverse function theorem as a generalization of the linear case] Let

$$F \; : \; \mathbb{R}^n \to \mathbb{R}^n \qquad F(x) = Ax,$$

where $A \; : \; n \times n$ is nonsingular. By simple linear algebra, the linear map F is (globally) bijective.

Note that

$$DF(p) = A$$

for any $p \in \mathbb{R}^n$.

Example 2 [Simple illustration] Consider the smooth map

$$F \; : \; \mathbb{R}^2 \to \mathbb{R}^2 \qquad F(x,y) = \left(x^2 + y^2, xy \right).$$

Then,

$$DF(1,0) = \begin{bmatrix} 2x & 2y \\ y & x \end{bmatrix}_{(x,y)=(1,0)} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

is non-singular, which means that F is a diffeomorphism near $(1,0)$.

Note that F is not a bijective map: $F(1,1) = F(-1,-1)$.

Theorem [Implicit Function Theorem] Let $W \subset \mathbb{R}^n \times \mathbb{R}^k$ be an open set and $F \; : \; W \to \mathbb{R}^k$,

$$(x,y) = \left(x^1, \ldots, x^n, y^1, \ldots, y^k \right) \quad \overset{F}{\longmapsto} \quad F(x,y) = \left(F^1(x,y), \ldots, F^k(x,y) \right)$$

a smooth map. Let $(p,q) = \left(p^1 \ldots, p^n, q^1, \ldots, q^k \right) \in W$ with $F(p,q) = 0$ and suppose that

$$D_y F(p,q) = \begin{pmatrix} \dfrac{\partial F^1}{\partial y^1}(p,q) & \dfrac{\partial F^1}{\partial y^2}(p,q) & \cdots & \dfrac{\partial F^1}{\partial y^k}(p,q) \\[2em] \dfrac{\partial F^2}{\partial y^1}(p,q) & \dfrac{\partial F^2}{\partial y^2}(p,q) & \cdots & \dfrac{\partial F^2}{\partial y^k}(p,q) \\[2em] \vdots & \vdots & \cdots & \vdots \\[2em] \dfrac{\partial F^k}{\partial y^1}(p,q) & \dfrac{\partial F^k}{\partial y^2}(p,q) & \cdots & \dfrac{\partial F^k}{\partial y^k}(p,q) \end{pmatrix}$$

is nonsingular. Then, there exist neighborhoods U_0 of p and V_0 of q and a smooth map $\Phi : U_0 \to V_0$ such that $U_0 \times V_0 \subset W$ and

$$(x, y) \in U_0 \times V_0, \ F(x, y) = 0 \quad \text{if and only if} \quad y = \Phi(x).$$

Furthermore, we have $D\Phi(p) = - \left(D_y F(p, q) \right)^{-1} D_x F(p, q)$, where

$$D_x F(p, q) = \begin{pmatrix} \dfrac{\partial F^1}{\partial x^1}(p, q) & \dfrac{\partial F^1}{\partial x^2}(p, q) & \cdots & \dfrac{\partial F^1}{\partial x^n}(p, q) \\[2ex] \dfrac{\partial F^2}{\partial x^1}(p, q) & \dfrac{\partial F^2}{\partial x^2}(p, q) & \cdots & \dfrac{\partial F^2}{\partial x^n}(p, q) \\[2ex] \vdots & \vdots & \cdots & \vdots \\[2ex] \dfrac{\partial F^k}{\partial x^1}(p, q) & \dfrac{\partial F^k}{\partial x^2}(p, q) & \cdots & \dfrac{\partial F^k}{\partial x^n}(p, q) \end{pmatrix}.$$

$$\{(x, y) \in W \ : \ F(x, y) = 0\}$$

3
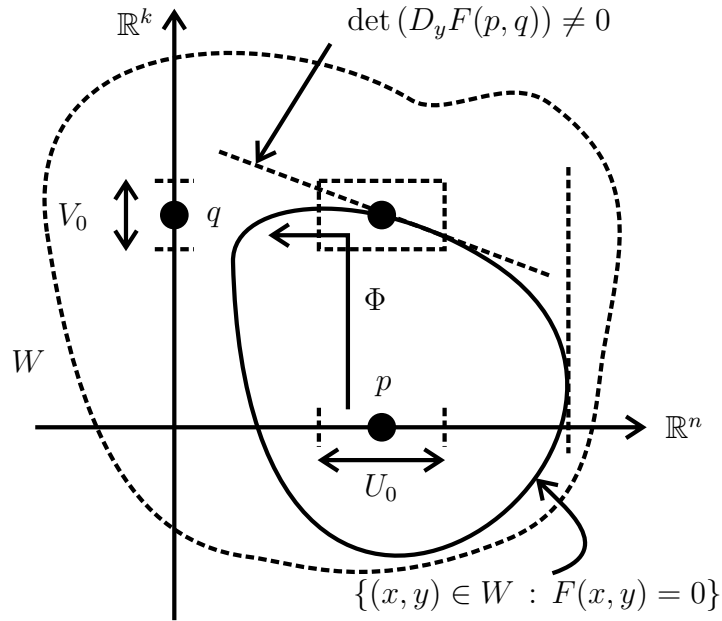
Example 1 [Implicit function theorem as a generalization of the linear case] Let

$$F \ : \ \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^k \qquad F\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

where $A \ : \ k \times n$ and the matrix $B \ : \ k \times k$ is nonsingular.

By simple linear algebra,

$$F(x, y) = 0 \quad \text{if and only if} \quad y = -B^{-1}Ax.$$

That is,

$$F(x, y) = 0 \quad \text{if and only if} \quad y = \Phi(x),$$

where

$$\Phi \ : \ \mathbb{R}^n \to \mathbb{R}^k \qquad \Phi(x) = -B^{-1}Ax.$$

Note that

$$B = D_y F(p, q) \quad \text{and} \quad D_x F(p, q) = A$$

for all $(p, q) \in \mathbb{R}^n \times \mathbb{R}^k$.

Example 2 [Simple eigenvalues are smooth] Let $X_0 \in \mathsf{M}(n, \mathbb{R})$ be a symmetric matrix and u_0 be an unit-norm eigenvector associated with the simple eigenvalue λ_0:

$$X_0 u_0 = \lambda_0 u_0 \quad \text{and} \quad u_0^T u_0 = 1.$$

Then, there exists a neighborhood $U_0 \subset \mathsf{M}(n, \mathbb{R})$ of X_0, a neighborhood $V_0 \subset \mathbb{R}^n \times \mathbb{R}$ of (u_0, λ_0) and a smooth map

$$\Phi \ : \ U_0 \to V_0 \qquad \Phi(X) = (u(X), \lambda(X)),$$

such that $u(X_0) = u_0$, $\lambda(X_0) = \lambda_0$, and

$$Xu(X) = \lambda(X)u(X) \quad \text{and} \quad u(X)^T u(X) = 1,$$

for all $X \in U_0$. The derivative of the map Φ at X_0 is given by

$$D\Phi(X_0) = \begin{bmatrix} u_0^T \otimes (\lambda_0 I_n - X_0)^+ \\ u_0^T \otimes u_0^T \end{bmatrix}.$$

**Example 3 [Signal processing application: asymptotic performance analysis]** Consider the data model

$$y[k] = \theta s[k] + w[k], \qquad k = 1, 2, \ldots, K,$$

where $y[k] = (y_1[k], y_2[k], \ldots, y_n[k]) \in \mathbb{R}^n$ is the observation vector at the kth time instant, K is the number of data samples available,

$$\theta \in \mathsf{S}^{n-1}_+(\mathbb{R}) = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n : \|x\| = 1 \text{ and } x_n > 0\}$$

is a deterministic parameter (the channel), $s[k] \in \mathbb{R}$ is a zero-mean, unit-power Gaussian random variable and $w[k] \in \mathbb{R}^n$ is a random vector distributed as $w \sim \mathcal{N}(0, \sigma^2 I_n)$. Here, $\mathcal{N}(\mu, \Sigma)$ stands for the normal (Gaussian) distribution with mean μ and covariance matrix Σ. We assume that the source $s[k]$ and the observation noise $w[k]$ are independent. Also, both are white random processes (independent from sample to sample).

You can view this data model as a simplified version of a multiple sensor system: an array of n sensors (or antennas) observe a Gaussian source s in a Gaussian background environment. Each data vector y represents an array snapshot. The parameter θ_i represents the gain of the channel established between the source s and the ith sensor.
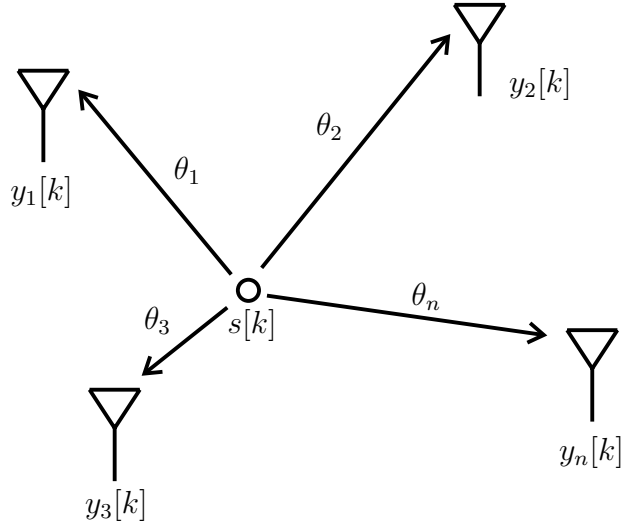
We assume the multi-channel vector $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$ is unknown and we want to estimate it from the available sample. Once θ is known

(estimated), we can return to the data and estimate the source s (by "reverting" the action of the channel).

The maximum-likelihood (ML) estimate of θ,

$$\widehat{\theta}_K = \underset{\theta \in \mathsf{S}_+^{n-1}(\mathbb{R})}{\mathsf{argmax}} \ p(y_1, \ldots, y_K; \theta),$$

is easily seen to be given by the unit-norm eigenvector (with last coordinate positive) which is associated with the maximum eigenvalue of the sample covariance matrix

$$\widehat{R}_K = \frac{1}{K} \sum_{k=1}^{K} y[k] y[k]^T.$$

Thus, $\widehat{\theta}_K = \phi(\widehat{R}_K)$, where ϕ stands for the map just described.

In this example, we are interested in evaluating the performance of this estimator. The performance index is the mean-square error (MSE) of the estimate

$$\mathsf{MSE} = \mathsf{E}\left\{ \left\| \widehat{\theta}_K - \theta \right\|^2 \right\}.$$

Since it is difficult to obtain the exact distribution of the statistic $\widehat{\theta}_K$, we resort to an asymptotic analysis. That is, we let K (the sample size) converge to infinity and obtain the limiting distribution of $\widehat{\theta}_K$. For any finite K, the asymptotic study provides us with an approximation for the MSE.

A fundamental tool in asymptotic analysis is the δ-method: let $x_K \in \mathbb{R}^n$ denote a sequence of random vectors satisfying

$$\sqrt{K} (x_K - \mu) \xrightarrow{d} \mathcal{N}(0, \Sigma),$$

where $\xrightarrow{d}$ stands for convergence in distribution (as $K \to +\infty$), and let $f : \mathbb{R}^n \to \mathbb{R}^m$ denote a map which is of class C^1 (continuous partial derivatives) in a neighborhood of μ. Then,

$$\sqrt{K} (f(x_K) - f(\mu)) \xrightarrow{d} \mathcal{N}\left(0, Df(\mu)\Sigma Df(\mu)^T\right),$$

where $Df(\mu)$ stands for the derivative of f at the point μ. In words, differentiable maps f transform asymptotically Gaussian random sequences x_K in asymptotically Gaussian random sequences $f(x_K)$.

In our context, it can be shown (trivial application of the central limit theorem) that

$$\sqrt{K}\left(\mathsf{vec}\left(\widehat{R}_K\right) - \mathsf{vec}(R)\right) \xrightarrow{d} \mathcal{N}\left(0, \Sigma\right),$$

for a certain covariance matrix Σ (known, but not explicited here) and where

$$R = \mathsf{E}\left\{y[k]y[k]^T\right\} = \theta\theta^T + \sigma^2 I_n$$

denotes the correlation matrix corresponding to our data model. Furthermore, note that the maximum eigenvalue of R is $\lambda_{\mathsf{max}} = 1 + \sigma^2$. The previous example has shown that ϕ is smooth in a neighborhood of R and its derivative is given by

$$D\phi(R) = \theta^T \otimes (\lambda_{\mathsf{max}} I_n - R)^+.$$

Thus, we have

$$\sqrt{K}\left(\widehat{\theta}_K - \theta\right) \xrightarrow{d} \mathcal{N}\left(0, D\phi(R)\Sigma D\phi(R)^T\right),$$

from which follows the approximation (for a given K)

$$\widehat{\theta}_K - \theta \sim \mathcal{N}\left(0, \frac{1}{K}D\phi(R)\Sigma D\phi(R)^T\right).$$

That is,

$$
\begin{aligned}
\mathsf{MSE} &= \mathsf{E}\left\{\left\|\widehat{\theta}_K - \theta\right\|^2\right\} \\
&= \mathsf{tr}\left(\mathsf{E}\left\{\left(\widehat{\theta}_K - \theta\right)\left(\widehat{\theta}_K - \theta\right)^T\right\}\right) \\
&\simeq \frac{1}{K}\mathsf{tr}\left(D\phi(R)\Sigma D\phi(R)^T\right).
\end{aligned}
$$

Figure 1 shows the results from a simulation example where: the observation vector y has dimension $n = 10$, $\theta = (1/\sqrt{n}, \ldots, 1/\sqrt{n})$, the signal-to-noise ratio

$$\mathsf{SNR} = \frac{\mathsf{E}\left\{\|\theta s\|^2\right\}}{\mathsf{E}\left\{\|w\|^2\right\}} = \frac{1}{n\sigma^2}$$

7

is fixed at 10 dB and the sample size K is varied between $K_{\mathsf{min}} = 10$ and $K_{\mathsf{max}} = 100$. As can be seen, the asymptotic analysis provides a
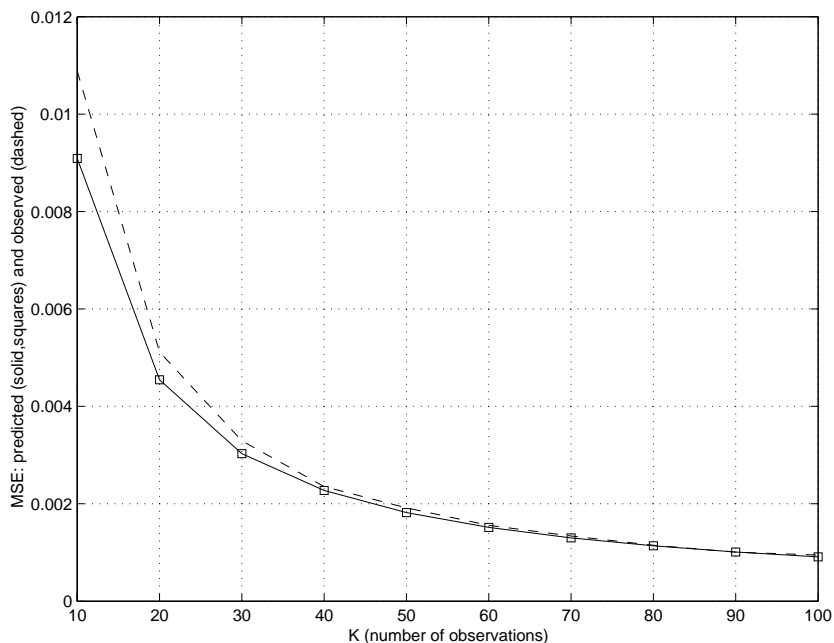

Figure 1: MSE of $\widehat{\theta}_K$: predicted (solid,squares) and observed (dashed)

good approximation of the MSE of the ML estimate $\widehat{\theta}_K$ (of course, the approximation gets tighter as $K \to \infty$).

Theorem [Rank theorem] Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ be open sets, and $F : U \to V$ a smooth map with constant rank k, that is, $\mathsf{rank}\,(DF(x)) = k$ for each $x \in U$. Let $p \in U$. Then, there exist neighborhoods $U_0 \subset U$ of p and $V_0 \subset V$ of $q = F(p)$ and diffeomorphisms $\varphi : U_0 \to \widehat{U}_0$ and $\psi : V_0 \to \widehat{V}_0$ such that
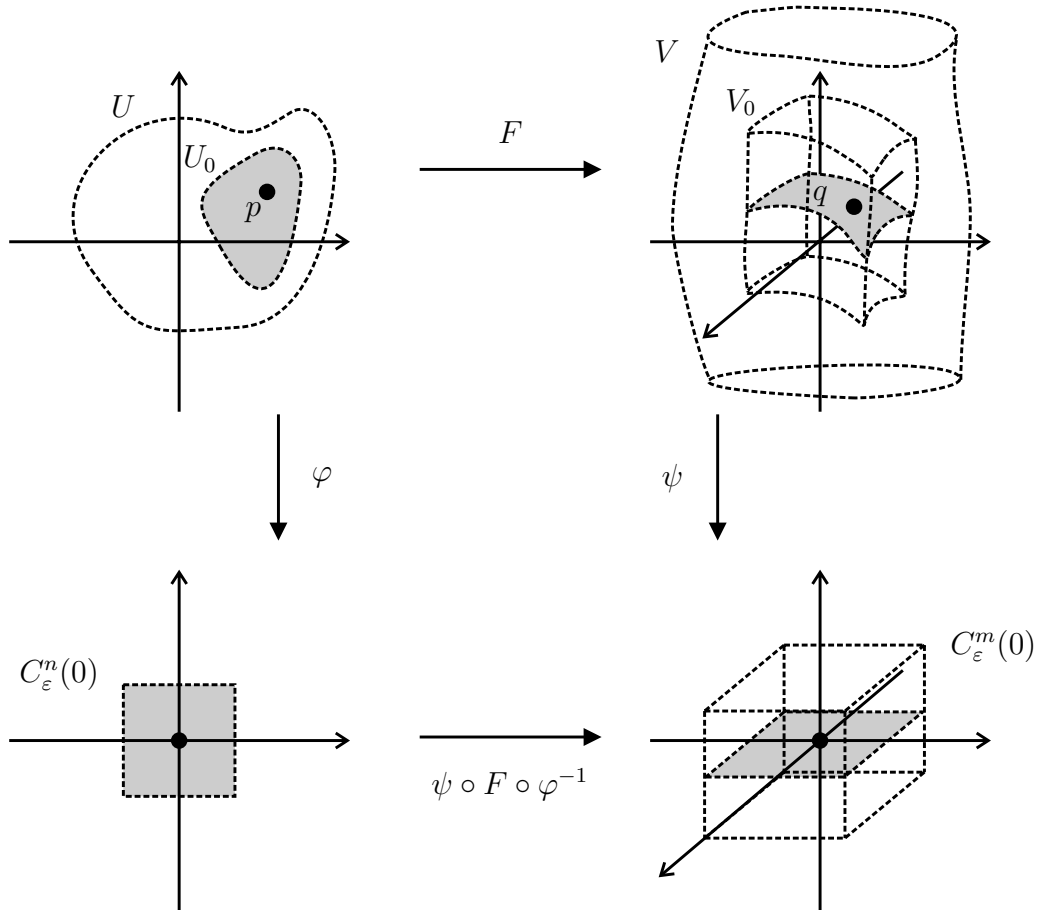
$$\psi \circ F \circ \varphi^{-1} \left(x^1, \ldots, x^k, x^{k+1}, \ldots, x^n \right) = \left(x^1, \ldots, x^k, 0, \ldots, 0 \right).$$

The neighborhoods U_0 and V_0 can be chosen such that: **(i)** $\widehat{U}_0 = C_\varepsilon^n(0)$ and $\widehat{V}_0 = C_\varepsilon^m(0)$ or **(ii)** $\widehat{U}_0 = B_\varepsilon^n(0)$ and $\widehat{V}_0 = B_\varepsilon^m(0)$, for any chosen $\varepsilon > 0$.

▷ *Intuition: looks like a nonlinear generalization of the SVD for linear*

8

Definition [Rank of a smooth map, immersions, submersions] Let $F : M \to N$ be a smooth map between smooth manifolds. The rank of F at $p \in M$ is the dimension of the linear subspace $\mathsf{Im}\, F_*(T_p M) \subset T_{F(p)} N$. Equivalently, it is the rank of the Jacobian matrix $\mathsf{rank}\, D\widehat{F}(\varphi(p))$ in any smooth chart.

We say that F has constant rank k if the rank of F at any $p \in M$ is k.

The smooth map $F : M \to N$ is called an immersion if F_* is injective at every point. Equivalently, if $\mathsf{rank}\, F = \mathsf{dim}\, M$ at every point.

The smooth map $F : M \to N$ is called a submersion if F_* is surjective

at every point. Equivalently, if $\mathsf{rank}\, F = \mathsf{dim}\, N$ at every point.

Example 1 [An immersion of the unit-sphere] Consider the map

$$F\ :\ \mathsf{S}^{n-1}(\mathbb{R}) \to \mathsf{M}(n, \mathbb{R}) \qquad F(u) = uu^T.$$

We already know that F is smooth. The map F is also an immersion.

Example 2 [A submersion onto the unit-sphere] Consider the map

$$F\ :\ \mathbb{R}^n - \{0\} \to \mathsf{S}^{n-1}(\mathbb{R}) \qquad F(x) = \frac{x}{\|x\|}.$$

We already know that F is smooth. The map F is also a submersion.

Example 3 [Product manifolds] Let M and N be smooth manifolds.

For fixed $q \in N$, the inclusion map

$$\iota_q\ :\ M \to M \times N \qquad \iota(p) = (p, q)$$

is an immersion.

The projection map

$$\pi_M\ :\ M \times N \to M \qquad \pi_M(p, q) = p$$

is a submersion.

Lemma [Composition of immersions and submersions] The composition of immersions is an immersion. The composition of submersions is a submersion.

Theorem [Inverse function theorem for manifolds] Let $F\ :\ M \to N$ be a smooth map between manifolds. Let $p \in M$ and suppose $F_*\ :\ T_pM \to T_{F(p)}N$ is an isomorphism (equivalently, a bijective linear map). Then there exist neighborhoods U_0 of p and V_0 of $F(p)$ such that $F\,|_{U_0}\ :\ U_0 \to V_0$ is a diffeomorphism.

▷ *Intuition: the bijectivity of F_* transpires <u>locally</u> to F*

$$F_* : T_pM \to T_{F(p)}N$$
(isomorphism)

T_pM

U_0

p •

M

$T_{F(p)}N$

V_0

$F(p)$ •

N

$$F : M \to N$$
(local diffeomorphism)
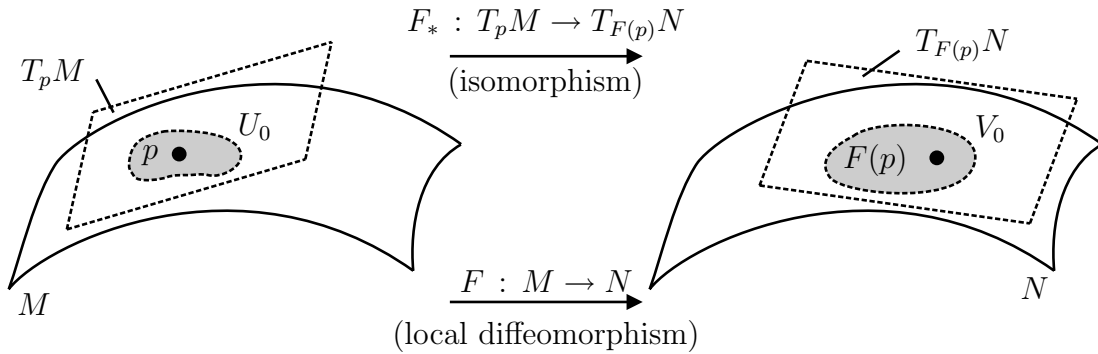
▷ *Remark that the inverse map $F^{-1} : V_0 \to U_0$ is smooth*

Example 1 [**Cholesky decomposition is a diffeomorphism**] The Cholesky decomposition asserts that for any $P \in \mathsf{P}(n, \mathbb{R})$ there is an unique $L \in \mathsf{L}^+(n, \mathbb{R})$ such that
$$P = LL^T.$$

Thus, we can define a map

$$\mathsf{Cholesky} : \mathsf{P}(n, \mathbb{R}) \to \mathsf{L}^+(n, \mathbb{R})$$

which, given a positive-definite matrix P, computes its Cholesky factor L such that $P = LL^T$.

The purpose of this example is to show that the map $\mathsf{Cholesky}$ is smooth.

We already know that the map

$$F : \mathsf{L}^+(n, \mathbb{R}) \to \mathsf{P}(n, \mathbb{R}) \qquad F(L) = LL^T$$

is bijective (linear algebra) and smooth. Remark that the map $\mathsf{Cholesky}$ is the inverse map of F.

Also, by exploiting the isomorphisms

$$T_{L_0}\mathsf{L}^+(n, \mathbb{R}) \simeq \mathsf{L}(n, \mathbb{R}) \quad \text{and} \quad T_{F(L_0)}\mathsf{P}(n, \mathbb{R}) \simeq \mathsf{S}(n, \mathbb{R}),$$

11

we have computed a representation of the push-forward map F_* : $T_{L_0}\mathsf{L}^+(n, \mathbb{R}) \to T_{F(L_0)}\mathsf{P}(n, \mathbb{R})$ as

$$F_* : \mathsf{L}(n, \mathbb{R}) \to \mathsf{S}(n, \mathbb{R}) \qquad F_*(\Delta) = \Delta L_0^T + L_0 \Delta^T.$$

Now, as soon as we show that F_* is an isomorphism, we can use the last theorem to conclude that $\mathsf{Cholesky} = F^{-1}$ is smooth (because it is smooth on a neighborhood of any given point $P_0 = F(L_0) \in \mathsf{P}(n, \mathbb{R})$).

To prove that the linear map F_* is bijective it suffices to prove that F_* is injective because $\dim \mathsf{L}(n, \mathbb{R}) = \dim \mathsf{S}(n, \mathbb{R})$ (thus, the surjectivity of F_* will follow automatically).

To prove that F_* is injective, we must show that $\mathsf{Ker}\, F_* = \{0\}$. So, let $\Delta \in \mathsf{L}(n, \mathbb{R})$ satisfy $F_*(\Delta) = 0$, that is,

$$\Delta L_0^T + L_0 \Delta^T = 0.$$

Pre-multiplying by L_0^{-1} and post-multiplying by $\left(L_0^T\right)^{-1}$ both sides of the equation yields

$$\left(L_0^{-1}\Delta\right) + \left(L_0^{-1}\Delta\right)^T = 0.$$

Note that L_0^{-1} is a lower-triangular matrix and $\Psi = L_0^{-1}\Delta$ also (product of two lower-triangular matrices). But,

$$\Psi + \Psi^T = 0 \quad \text{and} \quad \Psi : \text{lower-triangular} \quad \Rightarrow \quad \Psi = 0.$$

As a consequence, $\Delta = L_0\Psi = 0$.

We conclude that the map $\mathsf{Cholesky}$ is smooth. In fact, it is a diffeomorphism (because its inverse F is also smooth).

Theorem [Rank theorem for manifolds] Suppose that the smooth map $F : M \to N$ has constant rank k, with $\dim M = m$ and $\dim N = n$. Then, for any given $p \in M$, there exist smooth charts (U, φ) containing p and (V, ψ) containing $F(p)$ such that the coordinate representation $\widehat{F} = \psi \circ F \circ \varphi^{-1}$ is given by

$$\widehat{F}\left(x^1, x^2, \ldots, x^k, x^{k+1}, \ldots, x^m\right) = \left(x^1, x^2, \ldots, x^k, \underbrace{0, \ldots, 0}_{n-k \text{ zeros}}\right).$$
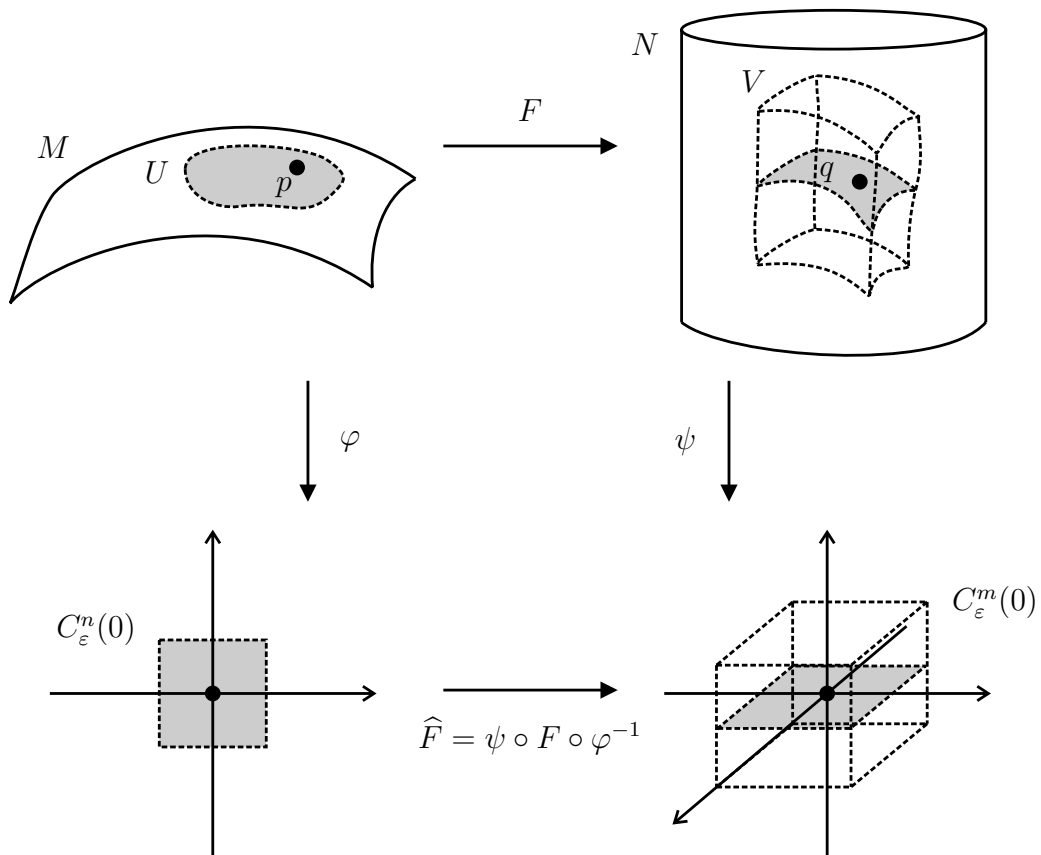
Theorem [Constant rank and immersions, submersions and diffeomorphisms] Let $F : M \to N$ be a smooth map of constant rank.

    (a) If F is injective, then it is an immersion

    (b) If F is surjective, then it is a submersion

    (c) If F is bijective, then it is a diffeomorphism

Example 1 [An immersion of the unit-circle] Consider the map

$$F : \mathsf{S}^1(\mathbb{R}) \to \mathsf{M}(2, \mathbb{R}) \qquad F(u) = \begin{bmatrix} u & Ju \end{bmatrix},$$

where

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

acts on vectors in $\mathbb{R}^2$ by rotating them by $\pi/2$ radians in the counter-clockwise direction.

The map F is smooth (why?). The goal of this example is to show that F is an immersion, without computing in coordinates.

Note that F is injective. If we prove that F has constant rank, we are done (see last theorem). Let $p, q \in \mathsf{S}^1(\mathbb{R})$. We must show that the two linear maps

$$F_{*p} : T_p\mathsf{S}^1(\mathbb{R}) \to T_{F(p)}\mathsf{M}(2, \mathbb{R}) \quad \text{and} \quad F_{*q} : T_q\mathsf{S}^1(\mathbb{R}) \to T_{F(q)}\mathsf{M}(2, \mathbb{R})$$

have the same rank.

The trick consists in noting that, for any fixed rotation $Q \in \mathsf{SO}(2)$, we have $F \circ L_Q = \widehat{L}_Q \circ F$ or, equivalently, the commutative diagram:

$$
\begin{array}{ccc}
\mathsf{S}^1(\mathbb{R}) & \xrightarrow{\;\;L_Q\;\;} & \mathsf{S}^1(\mathbb{R}) \\
\downarrow{\scriptstyle F} & & \downarrow{\scriptstyle F} \\
\mathsf{M}(2, \mathbb{R}) & \xrightarrow[\;\;\widehat{L}_Q\;\;]{} & \mathsf{M}(2, \mathbb{R})
\end{array}
$$

where

$$L_Q : \mathsf{S}^1(\mathbb{R}) \to \mathsf{S}^1(\mathbb{R}) \qquad L_Q(u) = Qu$$

and

$$\widehat{L}_Q : \mathsf{M}(2, \mathbb{R}) \to \mathsf{M}(2, \mathbb{R}) \qquad \widehat{L}_Q(X) = QX.$$

Note that both L_Q and $\widehat{L}_Q$ are smooth (why?). In fact they are diffeomorphisms because their inverse maps correspond to L_{Q^T} and $\widehat{L}_{Q^T}$, respectively, which are smooth.

Now, choose Q such that $L_Q(p) = q$. The previous diagram induces the next one, expressed in terms of push-forwards:

$$
\begin{array}{ccc}
T_pS^1(\mathbb{R}) & \xrightarrow{\;\;L_{Q*}\;\;} & T_qS^1(\mathbb{R}) \\
\Big\downarrow{\scriptstyle F_{*p}} & & \Big\downarrow{\scriptstyle F_{*q}} \\
T_{F(p)}\mathsf{M}(2,\mathbb{R}) & \xrightarrow[\;\;\widehat{L}_{Q*}\;\;]{} & T_{F(q)}\mathsf{M}(2,\mathbb{R})
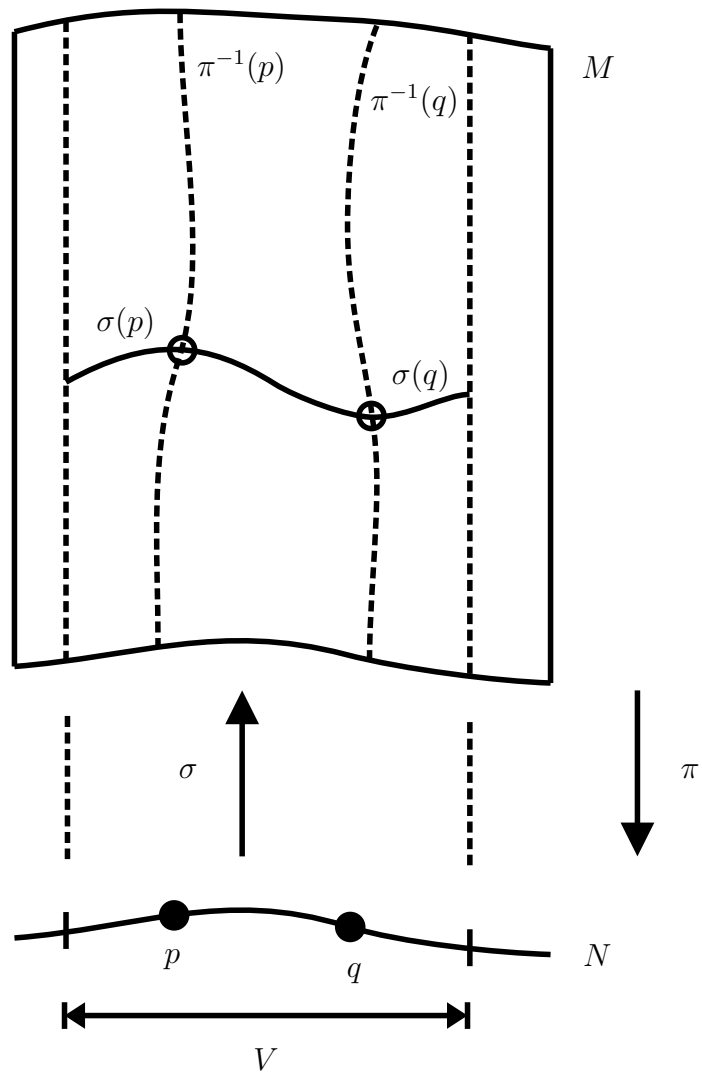\end{array}
$$

Equivalently:

$$
\widehat{L}_{Q*} \circ F_{*p} = F_{*q} \circ L_{Q*}.
$$

Since $\widehat{L}_{Q*}$ and L_{Q*} are isomorphisms,

$$
\mathsf{rank}\left(\widehat{L}_{Q*} \circ F_{*p}\right) = \mathsf{rank}\left(F_{*p}\right) \quad \text{and} \quad \mathsf{rank}\left(F_{*q} \circ L_{Q*}\right) = \mathsf{rank}\left(F_{*q}\right).
$$

The conclusion is $\mathsf{rank}\left(F_{*p}\right) = \mathsf{rank}\left(F_{*q}\right).$

Definition [Local section] Let $\pi : M \to N$ be a smooth map between smooth manifolds. A smooth local section of π is a pair (V, σ) where $V \subset N$ is open and $\sigma : V \to M$ is a smooth map satisfying $\pi \circ \sigma = \mathsf{id}_V$.

M

$\pi^{-1}(p)$

$\pi^{-1}(q)$

$\sigma(p)$

$\sigma(q)$

σ

π

N

p

q

V

▷ *Intuition: σ is a smooth choice of a representative in each fiber of π*

Lemma [Properties of submersions: part I] Let $\pi : M \to N$ be a smooth map between smooth manifolds. Suppose π is a submersion. Then, π is an open map. Moreover, for every $p \in M$, there exists a local section (V, σ) of π such that $p \in \sigma(V)$.

Lemma [Properties of submersions: part II] Let M, N, P be smooth manifolds and $\pi : M \to N$ be a surjective submersion. Then, a map $F : N \to P$ is smooth if and only if $\widehat{F} = F \circ \pi$ is smooth.

$$
\begin{array}{ccc}
M & & \\
{\scriptstyle \pi}\downarrow & \searrow{\scriptstyle \widehat{F}} & \\
N & \xrightarrow[F]{} & P
\end{array}
$$

$\triangleright$ *Intuition: smoothness of the "hard" map F can be investigated through the easier map $\widehat{F}$*

Example 1 [An immersion of the unit-sphere] Consider the map

$$F : \mathsf{S}^{n-1}(\mathbb{R}) \to \mathsf{M}(n, \mathbb{R}) \qquad F(u) = uu^T.$$

We already know that F is smooth. Here is an alternative proof of smoothness of F: use the facts that

$$\pi : \mathbb{R}^n - \{0\} \to \mathsf{S}^{n-1}(\mathbb{R}) \qquad \pi(x) = \frac{x}{\|x\|}$$

is a surjective submersion,
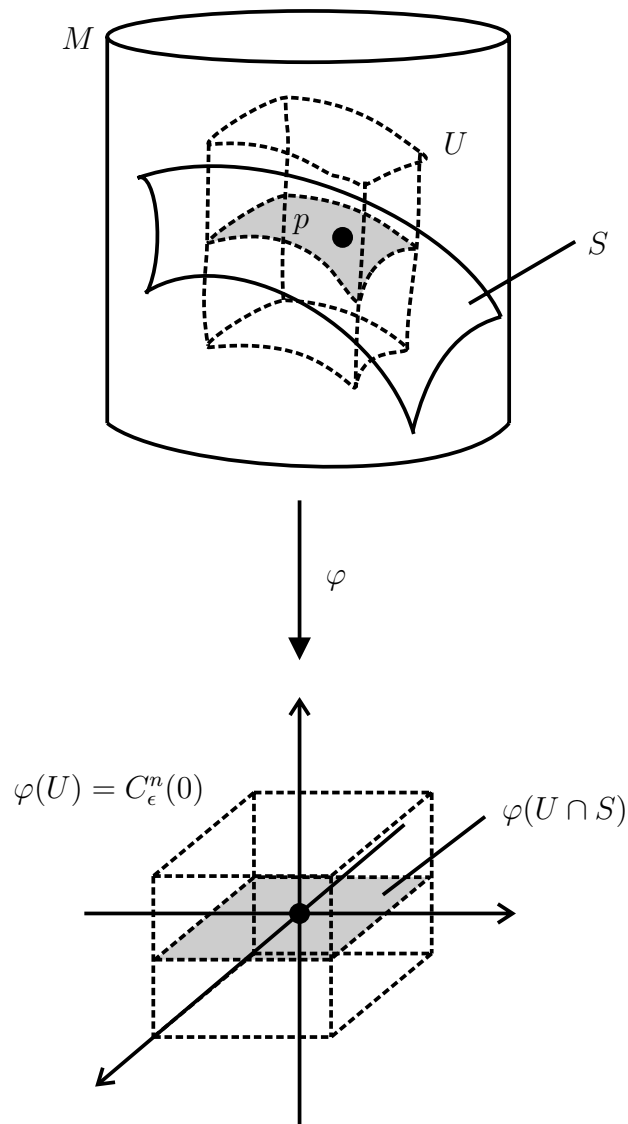
$$\widehat{F} : \mathbb{R}^n - \{0\} \to \mathsf{M}(n, \mathbb{R}) \qquad \widehat{F}(x) = \frac{xx^T}{\|x\|^2}.$$

is <u>clearly</u> smooth and $\widehat{F} = F \circ \pi$.

Definition [Embedded submanifold] Let M be an n-dimensional smooth manifold. A subset $S \subset M$ is called an embedded k-submanifold of M if,

for each point $p \in S$, there is a smooth chart (U, φ) centered at p with $\varphi(U) = C_\epsilon^n(0)$ and

$$\varphi(U \cap S) = \{(x^1, x^2, \ldots, x^k, x^{k+1}, \ldots, x^n) : x^{k+1} = x^{k+2} = \cdots = x^n = 0\}.$$

M

U

p

S

φ

$\varphi(U) = C_\epsilon^n(0)$

$\varphi(U \cap S)$

▷ *Intuition: the subset $S \subset M$ can be flattened (locally)*

▷ *Intuition: generalization of a linear subspace S of a vector space M*
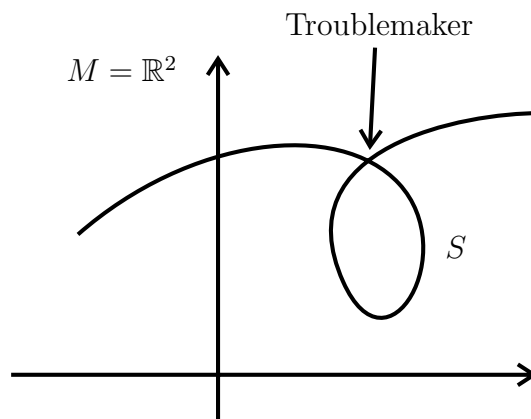
S is not an embedded submanifold of $\mathbb{R}^2$

Example 1 [Linear subspaces] If $S \subset \mathbb{R}^n$ is a linear subspace with (linear) dimension k. Then, S is an embedded k-submanifold of $\mathbb{R}^n$.

As an application, the linear subspace of symmetric matrices

$$\mathsf{S}(n, \mathbb{R}) = \{X \in \mathsf{M}(n, \mathbb{R}) \,:\, X = X^T\}$$

is an embedded $n(n + 1)/2$-submanifold of $\mathsf{M}(n, \mathbb{R})$.

Similarly, the linear subspace of skew-symmetric matrices

$$\mathsf{K}(n, \mathbb{R}) = \{X \in \mathsf{M}(n, \mathbb{R}) \,:\, X = -X^T\}$$

is an embedded $n(n - 1)/2$-submanifold of $\mathsf{M}(n, \mathbb{R})$.

Definition [Embedding submanifolds are local constructions] Let M be a smooth manifold. The subset $S \subset M$ is a embedded submanifold of M if and only if each $p \in S$ has a neighborhood $U \subset M$ such that $S \cap U$ is an embedded submanifold of U.

Lemma [Open subsets are embedded submanifolds] Let $U \subset M$ be an open subset of the n-dimensional smooth manifold M. Then, U is an embedded n-submanifold of M.

Example 1 [Positive definite matrices] The set of positive definite matrices

$$\mathsf{P}(n, \mathbb{R}) = \{X \in \mathsf{S}(n, \mathbb{R}) \, : \, X \succ 0\}$$

is an embedded $n(n+1)/2$-submanifold of $\mathsf{S}(n, \mathbb{R})$.

Definition [Embedding] A smooth map $F \, : \, N \to M$ between smooth manifolds is said to be an embedding if it is an immersion and a topological embedding (a homeomorphism of N onto its image $\widetilde{N} = F(N)$, viewed as a subspace of M).

Lemma [Useful criterion for detecting embeddings] Let the smooth map $F \, : \, M \to N$ be an injective immersion. If M is compact, F is an embedding.

Example 1 [An embedding of the unit-circle] Consider the map

$$F \, : \, \mathsf{S}^1(\mathbb{R}) \to \mathsf{M}(2, \mathbb{R}) \qquad F(u) = \begin{bmatrix} u & Ju \end{bmatrix},$$

where

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

We already know that F is a smooth immersion. Since $\mathsf{S}^1(\mathbb{R})$ is compact, F is an embedding.

Theorem [Embedded submanifolds are smooth manifolds] Let the subset $S \subset M$ be an embedded k-dimensional submanifold of M, where $\dim M = n$.

Then, as a subspace of M, S is a topological manifold of dimension k and it has an unique smooth structure such that the inclusion map $\iota \, : \, S \to M$ is a smooth embedding.

With this smooth structure on S, let (U, φ) be a smooth chart in M with $\varphi(U) = C_\epsilon^n(0)$ and

$$\varphi(U \cap S) = \{(x^1, x^2, \ldots, x^k, x^{k+1}, \ldots, x^n) \, : \, x^{k+1} = x^{k+2} = \cdots = x^n = 0\}.$$

Then, $(S \cap U, \widehat{\pi} \circ \varphi)$ is a smooth chart in S, where

$$\widehat{\pi}(x^1, \ldots, x^k, x^{k+1}, \ldots, x^n) = (x^1, \ldots, x^k).$$

Theorem [Smooth embeddings provide embedded submanifolds]
The image of a smooth embedding is an embedded submanifold.

Example 1 [$\mathsf{SO}(2)$ is an embedded submanifold of $\mathsf{M}(2, \mathbb{R})$] The subset

$$\mathsf{SO}(2) = \left\{ \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \right\}$$

is an embedded submanifold of $\mathsf{M}(2, \mathbb{R})$ because $\mathsf{SO}(2) = F(\mathsf{S}^1(\mathbb{R}))$ where F is the embedding

$$F : \mathsf{S}^1(\mathbb{R}) \to \mathsf{M}(2, \mathbb{R}) \qquad F(u) = \begin{bmatrix} u & Ju \end{bmatrix}, \qquad J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Lemma [Composition of embeddings] The composition of embeddings is an embedding.

Theorem [Constant-rank level set theorem] Let $F : M \to N$ be a smooth map with constant rank k. Then, for each $c \in \mathsf{Im}\, F$, the level set

$$F^{-1}(c) = \{p \in M \ : \ F(p) = c\}$$

is a closed, embedded submanifold of dimension $\mathsf{dim}\, M - k$.

Example 1 [Stiefel] Let

$$\mathsf{O}(n, m) = \{X \in \mathsf{M}(n, m, \mathbb{R}) \ : \ X^T X = I_m\}$$

be the set of $n \times m$ orthonormal frames in $\mathbb{R}^n$.

The goal of this example is to show that $\mathsf{O}(n,m)$ is an embedded submanifold of $\mathsf{M}(n,m,\mathbb{R})$. The manifold $\mathsf{O}(n,m)$ is known as the Stiefel manifold.

Note that $\mathsf{O}(n,m) \subset \mathsf{F}(n,m,\mathbb{R})$ where

$$
\begin{aligned}
\mathsf{F}(n,m,\mathbb{R}) &= \left\{ X \in \mathsf{M}(n,m,\mathbb{R}) \,:\, \det(X^T X) \neq 0 \right\} \\
&= \left\{ X \in \mathsf{M}(n,m,\mathbb{R}) \,:\, \mathsf{rank}(X) = m \right\}
\end{aligned}
$$

is the set of $n \times m$ frames in $\mathbb{R}^n$. Remark that $\mathsf{F}(n,m,\mathbb{R})$ is an open subset of $\mathsf{M}(n,m,\mathbb{R})$.

We will show that $\mathsf{O}(n,m)$ is an embedded submanifold of $\mathsf{F}(n,m,\mathbb{R})$ and, since $\mathsf{F}(n,m,\mathbb{R})$ is embedded in $\mathsf{M}(n,m,\mathbb{R})$, it will follow that $\mathsf{O}(n,m)$ is embedded in $\mathsf{M}(n,m,\mathbb{R})$.

The prove that $\mathsf{O}(n,m)$ is embedded in $\mathsf{F}(n,m,\mathbb{R})$ we start by noting that $\mathsf{O}(n,m)$ is a level set of the smooth map

$$
F \,:\, \mathsf{F}(n,m,\mathbb{R}) \to \mathsf{M}(m,\mathbb{R}) \qquad F(X) = X^T X.
$$

Indeed, $\mathsf{O}(n,m) = F^{-1}(I_m)$.

If we show that F has constant rank, we are done. The rest of the proof establishes this.

Choose $A \in \mathsf{F}(n,m,\mathbb{R})$ arbitrarily. We have the isomorphisms

$$
T_A \mathsf{F}(n,m,\mathbb{R}) \simeq \mathsf{M}(n,m,\mathbb{R}) \quad \text{and} \quad T_{F(A)} \mathsf{M}(n,\mathbb{R}) \simeq \mathsf{M}(n,\mathbb{R}).
$$

In terms of this isomorphisms, the push-forward $F_{*A} \,:\, T_A \mathsf{F}(n,m,\mathbb{R}) \to T_{F(A)} \mathsf{M}(m,\mathbb{R})$ is represented by the linear map

$$
F_{*A} \,:\, \mathsf{M}(n,m,\mathbb{R}) \to \mathsf{M}(m,\mathbb{R}) \qquad F_{*A}(\Delta) = \Delta^T A + A^T \Delta.
$$

Note that the rank of F at A is given by

$$
\mathsf{rank}\, F_{*A} = nm - \dim \mathsf{Ker}\, F_{*A}.
$$

Choose another $B \in \mathsf{F}(n,m,\mathbb{R})$. Similarly, the push-forward at B is represented by the linear map

$$
F_{*B} \,:\, \mathsf{M}(n,m,\mathbb{R}) \to \mathsf{M}(m,\mathbb{R}) \qquad F_{*B}(\Delta) = \Delta^T B + B^T \Delta
$$

and we have
$$\mathsf{rank}\, F_{*B} = nm - \mathsf{dim}\, \mathsf{Ker}\, F_{*B}.$$

The next step is to prove that $\mathsf{dim}\, \mathsf{Ker}\, F_{*A} \leq \mathsf{dim}\, \mathsf{Ker}\, F_{*B}$. Let $G \in \mathsf{GL}(n, \mathbb{R})$ be chosen such that $B = G^{-T}A$ and consider the linear map

$$L \,:\, \mathsf{M}(n, m, \mathbb{R}) \to \mathsf{M}(n, m, \mathbb{R}) \qquad L(\Delta) = G\Delta.$$

Then,
$$L(\mathsf{Ker}\, F_{*A}) \subset \mathsf{Ker}\, F_{*B}.$$

Indeed, if $\Delta \in \mathsf{Ker}\, F_{*A}$. Then, $L(\Delta) \in \mathsf{Ker}\, F_{*B}$ because

$$
\begin{aligned}
F_{*B}\,(L(\Delta)) &= L(\Delta)^T B + B^T L(\Delta) \\
&= (G\Delta)^T B + B^T (G\Delta) \\
&= \Delta^T (G^T B) + (B^T G)\Delta \\
&= \Delta^T A + A^T \Delta \\
&= 0.
\end{aligned}
$$

Since L is injective, this shows that $\mathsf{dim}\, \mathsf{Ker}\, F_{*A} \leq \mathsf{dim}\, \mathsf{Ker}\, F_{*B}$.

Reversing the roles $A \rightleftarrows B$ yields the opposite inequality $\mathsf{dim}\, \mathsf{Ker}\, F_{*B} \leq \mathsf{dim}\, \mathsf{Ker}\, F_{*A}$.

We conclude that F has constant rank:

$$\mathsf{rank}\, F_{*A} = nm - \mathsf{dim}\, \mathsf{Ker}\, F_{*A} = nm - \mathsf{dim}\, \mathsf{Ker}\, F_{*B} = \mathsf{rank}\, F_{*B}.$$

To actually compute the rank of F we pick a nice point (lots of zeros in it):

$$A = \begin{bmatrix} I_m \\ 0 \end{bmatrix}.$$

Then,

$$
\begin{aligned}
\mathsf{Ker}\, F_{*A} &= \left\{ \begin{bmatrix} \Delta_1 \\ \Delta_2 \end{bmatrix} \,:\, \Delta_1^T + \Delta_1 = 0 \right\} \\
&= \left\{ \begin{bmatrix} \Delta_1 \\ \Delta_2 \end{bmatrix} \,:\, \Delta_1 \in \mathsf{K}(m, \mathbb{R}),\, \Delta_2 \in \mathsf{M}(n - m, m, \mathbb{R}) \right\}.
\end{aligned}
$$

We conclude that

$$
\begin{aligned}
\dim \mathsf{Ker}\, F_{A*} &= \dim \mathsf{K}(m, \mathbb{R}) + \dim \mathsf{M}(n - m, m, \mathbb{R}) \\
&= \frac{m(m - 1)}{2} + (n - m)m.
\end{aligned}
$$

Thus, $\mathsf{O}(n, m)$ is an embedded submanifold of $\mathsf{M}(n, m, \mathbb{R})$ and

$$
\dim \mathsf{O}(n, m) = \frac{m(m - 1)}{2} + (n - m)m.
$$

Special cases:

- (Unit-sphere) For $\mathsf{S}^{n-1}(\mathbb{R}) = \mathsf{O}(n, 1)$, we have

$$
\dim \mathsf{S}^{n-1}(\mathbb{R}) = n - 1
$$

- (Orthogonal group) For $\mathsf{O}(n) = \mathsf{O}(n, n)$, we have

$$
\dim \mathsf{O}(n) = \frac{n(n - 1)}{2}.
$$

Example 2 [Special orthogonal group $\mathsf{SO}(n)$] Since $\mathsf{SO}(n)$ is an open subset of the smooth manifold $\mathsf{O}(n)$, it is an embedded submanifold of $\mathsf{O}(n)$ and $\dim \mathsf{SO}(n) = \dim \mathsf{O}(n) = n(n - 1)/2$.

Since $\mathsf{O}(n)$ is embedded in $\mathsf{M}(n, \mathbb{R})$, we conclude that $\mathsf{SO}(n)$ is an embedded submanifold of $\mathsf{M}(n, \mathbb{R})$.

Example 3 [Matrices with fixed rank] Consider the subset of $n \times m$ matrices with rank k

$$
\mathsf{Rank}_{=k}(n, m, \mathbb{R}) = \{X \in \mathsf{M}(n, m, \mathbb{R}) : \mathsf{rank}\, X = k\}.
$$

To simplify notation we set $R = \mathsf{Rank}_{=k}(n, m, \mathbb{R})$. We show that R is an embedded submanifold of $\mathsf{M}(n, m, \mathbb{R})$.

Let

$$
X_0 = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in R \quad \text{where} \quad A : k \times k.
$$

Without loss of generality we assume $\det(A) \neq 0$ (the general case $X_0 \in R$ can be reverted into this canonical format through pre- and post-multiplication by suitable permutation matrices).

Define the set

$$U = \left\{ \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} : \det(X_{11}) \neq 0 \text{ where } X_{11} : k \times k \right\}.$$

Note that U is an open subset containing X_0. We show that $U \cap R$ is embedded in U.

This is established by noting that

$$\begin{aligned}
X \in U \cap R \quad &\Leftrightarrow \quad \mathsf{rank}\left(\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \right) = k, \ \det(X_{11}) \neq 0 \\
&\Leftrightarrow \quad \mathsf{rank}\left(\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} I & -X_{11}^{-1}X_{12} \\ 0 & I \end{bmatrix} \right) = k, \ \det(X_{11}) \neq 0 \\
&\Leftrightarrow \quad \mathsf{rank}\left(\begin{bmatrix} X_{11} & 0 \\ X_{21} & X_{22} - X_{21}X_{11}^{-1}X_{12} \end{bmatrix} \right), \ \det(X_{11}) \neq 0 \\
&\Leftrightarrow \quad X_{22} - X_{21}X_{11}^{-1}X_{12} = 0, \ \det(X_{11}) \neq 0.
\end{aligned}$$

This means that $U \cap R = F^{-1}(0)$ where

$$F : U \to \mathsf{M}(n-k, m-k, \mathbb{R}) \qquad F(X) = X_{22} - X_{21}X_{11}^{-1}X_{12}.$$

We now show that F has constant rank. In fact, we show that F is a submersion.

We have

$$F_* : T_{X_0}U \simeq \mathsf{M}(n, m, \mathbb{R}) \to T_{F(X_0)}\mathsf{M}(n-k, m-k, \mathbb{R}) \simeq \mathsf{M}(n-k, m-k, \mathbb{R})$$

with

$$F_*\left(\begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{bmatrix} \right) = \Delta_{22} + \text{ map depending linearly on } \Delta_{11}, \Delta_{12}, \Delta_{21}.$$

Thus, for any given $\Psi \in \mathsf{M}(n-k, m-k, \mathbb{R})$ we have

$$F_*\left(\begin{bmatrix} 0 & 0 \\ 0 & \Psi \end{bmatrix} \right) = \Psi.$$

This shows that F_* is surjective (that is, F is a submersion).

Lemma [Identifications for tangent spaces] Let $F : M \to N$ be a smooth map with constant rank k. Let $c \in \mathsf{Im}\, F$. Thus, the level set $S = F^{-1}(c)$ is an embedded submanifold of M and the dimension of S is $d = \dim M - k$.
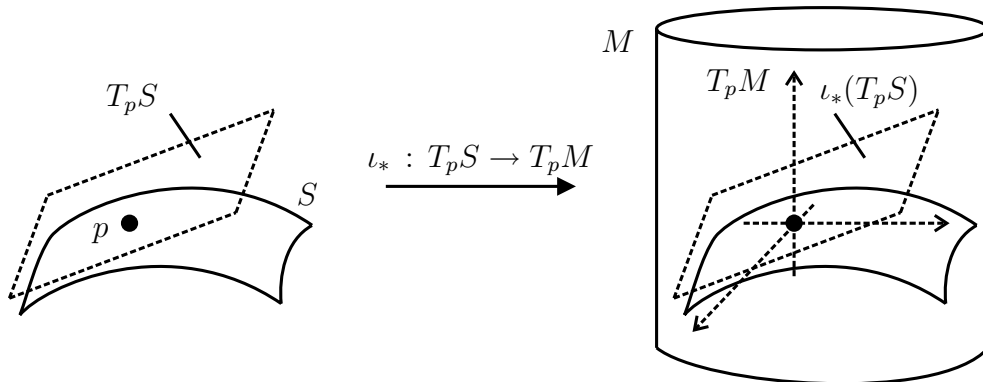
Since the inclusion $\iota : S \to M$ is an embedding (in particular, an immersion), it follows that, for any $p \in S$, the push-forward $\iota_* : T_pS \to T_PM$ is injective and

$$\iota_*(T_pS) \subset T_pM$$

is a d-dimensional subspace of T_pM. We usually make the identification $T_pS \simeq \iota_*(T_pS)$.

In the current setup, we have $\iota_* (T_pS) = \mathsf{Ker}\, F_{*p}$. Thus,

$$T_pS \simeq \mathsf{Ker}\, F_{*p}.$$

$\triangleright$ *T_pS after being push-forwarded by ι_* appears as a subspace of T_pM*

Example 1 [Unit-sphere] The unit-sphere $\mathsf{S}^{n-1}(\mathbb{R})$ is a level set of the constant-rank map

$$F : \mathbb{R}^n - \{0\} \to \mathbb{R} \qquad F(x) = x^T x.$$

Indeed, $\mathsf{S}^{n-1}(\mathbb{R}) = F^{-1}(1)$. Thus, we have

$$T_p\mathsf{S}^{n-1}(\mathbb{R}) \simeq \mathsf{Ker}\, F_{*p}$$

for any $p \in S^{n-1}(\mathbb{R})$.

Using the identifications

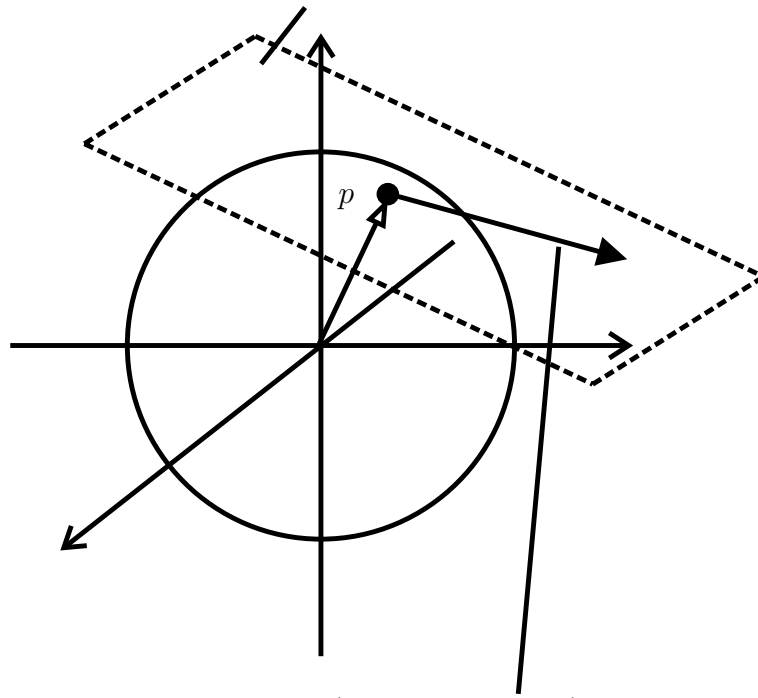$$T_p \mathbb{R}^n - \{0\} \simeq \mathbb{R}^n \quad \text{and} \quad T_{F(p)} \mathbb{R} \simeq \mathbb{R}$$

the push-forward $F_{*p} : T_p \mathbb{R}^n - \{0\} \to T_{F(p)} \mathbb{R}$ is represented by the linear map

$$F_{*p} : \mathbb{R}^n \to \mathbb{R} \qquad F_{*p}(\delta) = \delta^T p + p^T \delta.$$

Hence,

$$T_p S^{n-1}(\mathbb{R}) \simeq \mathsf{Ker}\, F_{*p} = \{\delta \in \mathbb{R}^n : p^T \delta = 0\}.$$

$$T_p S^2(\mathbb{R}) \simeq \iota_* \left(T_p S^2(\mathbb{R}) \right) \subset T_p M$$

$$\delta = (\delta^1, \ldots, \delta^n) \in \mathbb{R}^n \simeq \delta^1 \frac{\partial}{\partial x^1}\bigg|_p + \cdots + \delta^n \frac{\partial}{\partial x^n}\bigg|_p \in T_p M$$

Example 2 [Orthogonal group] The orthogonal group $\mathsf{O}(n)$ is a level set of the constant-rank map

$$F : \mathsf{GL}(n, \mathbb{R}) \to \mathsf{M}(n, \mathbb{R}) \qquad F(X) = X^T X.$$

28

Indeed, $\mathsf{O}(n) = F^{-1}(I_n)$. Thus, we have

$$T_Q\mathsf{O}(n) \simeq \mathsf{Ker}\, F_{*Q}$$

for any $Q \in \mathsf{O}(n)$.

Using the identifications

$$T_Q\mathsf{GL}(n, \mathbb{R}) \simeq \mathsf{M}(n, \mathbb{R}) \quad \text{and} \quad T_{F(Q)}\mathsf{M}(n, \mathbb{R}) \simeq \mathsf{M}(n, \mathbb{R})$$

the push-forward $F_{*Q} \,:\, T_Q\mathsf{GL}(n, \mathbb{R}) \to T_{F(Q)}\mathsf{M}(n, \mathbb{R})$ is represented by the linear map

$$F_{*Q} \,:\, \mathsf{M}(n, \mathbb{R}) \to \mathsf{M}(n, \mathbb{R}) \qquad F_{*Q}(\Delta) = \Delta^T Q + Q^T \Delta.$$

Thus,
$$T_Q\mathsf{O}(n) \simeq \mathsf{Ker}\, F_{*Q} = \{QK \,:\, K \in \mathsf{K}(n, \mathbb{R})\},$$

where
$$\mathsf{K}(n, \mathbb{R}) = \{X \in \mathsf{M}(n, \mathbb{R}) \,:\, X = -X^T\}$$

is the set of $n \times n$ skew-symmetric matrices.

We also use the compact notation $T_Q\mathsf{O}(n) \simeq Q\mathsf{K}(n, \mathbb{R})$.

Proposition [Restricting the domain and/or range of smooth maps]
Let $F \,:\, M \to N$ be a smooth map.

If A is an embedded submanifold of M, then the map

$$F|_A \,:\, A \to N \qquad F|_A(p) = F(p)$$

is smooth.

If B is an embedded submanifold of N and $F(M) \subset B$, then the map

$$F|^B \,:\, M \to B \qquad F|^B(p) = F(p)$$

is smooth.

Example 1 [An immersion of the unit-sphere] Consider the map

$$F : \mathsf{S}^{n-1}(\mathbb{R}) \to \mathsf{M}(n, \mathbb{R}) \qquad F(u) = uu^T.$$

The map F is smooth because:

Step 1:
$$\widehat{F} : \mathbb{R}^n \to \mathsf{M}(n, \mathbb{R}), \qquad \widehat{F}(x) = xx^T$$

is <u>clearly</u> smooth

Step 2: $\mathsf{S}^{n-1}(\mathbb{R})$ is an embedded submanifold of $\mathbb{R}^n$

Step 3: $F = \widehat{F}|_{\mathsf{S}^{n-1}(\mathbb{R})}$

Example 2 [A submersion onto the unit-sphere] Consider the map

$$F : \mathbb{R}^n - \{0\} \to \mathsf{S}^{n-1}(\mathbb{R}) \qquad F(x) = \frac{x}{\|x\|}.$$

The map F is smooth because:

Step 1:
$$\widehat{F} : \mathbb{R}^n - \{0\} \to \mathbb{R}^n, \qquad \widehat{F}(x) = \frac{x}{\|x\|}$$

is <u>clearly</u> smooth

Step 2: $\mathsf{S}^{n-1}(\mathbb{R})$ is an embedded submanifold of $\mathbb{R}^n$

Step 3: $F = \widehat{F}|^{\mathsf{S}^{n-1}(\mathbb{R})}$

Example 3 [Concatenating the techniques] Consider the map

$$F : \mathsf{O}(n) \to \mathsf{S}^{n-1}(\mathbb{R}) \qquad F(X) = F([\, x_1 \, x_2 \, \cdots \, x_n \,]) = x_1.$$

The map F is smooth because:

Step 1:
$$\widehat{F} : \mathsf{M}(n, \mathbb{R}) \to \mathbb{R}^n, \qquad \widehat{F}(X) = F([\, x_1 \, x_2 \, \cdots \, x_n \,]) = x_1$$

is <u>clearly</u> smooth

Step 2: $\mathsf{O}(n)$ is an embedded submanifold of $\mathsf{M}(n, \mathbb{R})$, hence,

$$\widehat{F}|_{\mathsf{O}(n)} : \mathsf{O}(n) \to \mathbb{R}^n, \qquad \widehat{F}|_{\mathsf{O}(n)}(X) = x_1$$

$$\mathsf{M}(n, \mathbb{R}) \xrightarrow{\;\widehat{F}\;} \mathbb{R}^n$$

$$\iota_{\mathsf{O}(n)} \uparrow \qquad \widehat{F}|_{\mathsf{O}(n)}$$

$$\mathsf{O}(n)$$

is smooth

Step 3: $\mathsf{S}^{n-1}(\mathbb{R})$ is an embedded submanifold of $\mathbb{R}^n$ and $\widehat{F}|_{\mathsf{O}(n)}\,(\mathsf{O}(n)) \subset \mathsf{S}^{n-1}(\mathbb{R})$; hence,

$$F = \widehat{F}\big|_{\mathsf{O}(n)}^{\mathsf{S}^{n-1}(\mathbb{R})}$$

is smooth

$$\mathbb{R}^n$$

$$\widehat{F}|_{\mathsf{O}(n)} \qquad \iota_{\mathsf{S}^{n-1}(\mathbb{R})}$$

$$\mathsf{O}(n) \xrightarrow{\;F\;} \mathsf{S}^{n-1}(\mathbb{R})$$

Example 4 [Using identifications for computations] Let $F : A \to B$ be a smooth map between smooth manifolds. Assume that A and B are embedded in M and N, respectively. Suppose that there exists a smooth map $\widehat{F} : M \to N$ such that the following diagram commutes (i.e., $\iota_B \circ F = \widehat{F} \circ \iota_A$)

$$M \xrightarrow{\;\widehat{F}\;} N$$

$$\iota_A \uparrow \qquad\qquad \uparrow \iota_B$$

$$A \xrightarrow{\;F\;} B$$

For any $p \in A$, we have the corresponding diagram in terms of the push-forwards

$$\iota_{A*}(T_pA) \subset T_pM \xrightarrow{\widehat{F}_*} \iota_{B*}(T_{F(p)}B) \subset T_{F(p)}N$$

$$\Big\uparrow {\scriptstyle \iota_{A*}} \qquad\qquad \Big\uparrow {\scriptstyle \iota_{B*}}$$

$$T_pA \xrightarrow{\;F_*\;} T_{F(p)}B$$

This means that we can represent the push-forward map $F_* : T_pA \to T_{F(p)}B$ by the push-forward map

$$\widehat{F}_* \; : \; \iota_{A*}(T_pA) \to \iota_{B*}(T_{F(p)}B).$$

Example 5 [An immersion of the unit-sphere] Consider the map

$$F \; : \; \mathsf{S}^{n-1}(\mathbb{R}) \to \mathsf{M}(n, \mathbb{R}) \qquad F(x) = xx^T.$$

We have the commutative diagram

$$\mathbb{R}^n \xrightarrow{\widehat{F}} \mathsf{M}(n, \mathbb{R})$$

$$\scriptstyle \iota_{\mathsf{S}^{n-1}(\mathbb{R})} \Big\uparrow \qquad \nearrow {\scriptstyle F}$$

$$\mathsf{S}^{n-1}(\mathbb{R})$$

It is easy to obtain the push-forward of $\widehat{F}$ at any point $p \in \mathbb{R}^n$:

$$\widehat{F}_* \; : \; T_p\mathbb{R}^n \simeq \mathbb{R}^n \to T_{\widehat{F}(p)}\mathsf{M}(n, \mathbb{R}) \simeq \mathsf{M}(n, \mathbb{R}) \qquad \widehat{F}(\delta) = \delta p^T + p\delta^T.$$

On the other hand,

$$T_p\mathsf{S}^{n-1}(\mathbb{R}) \simeq \{\delta \in \mathbb{R}^n \; : \; p^T\delta = 0\}.$$

In conclusion, we can represent the push-forward $F_* \; : \; T_p\mathsf{S}^{n-1}(\mathbb{R}) \to T_{F(p)}\mathsf{M}(n, \mathbb{R})$ by the linear map

$$F_* \; : \; \{\delta \; : \; p^T\delta = 0\} \to \mathsf{M}(n, \mathbb{R}) \qquad F_*(\delta) = \delta p^T + p\delta^T.$$

32

As an example, we can exploit the representation above to prove that the smooth map F is an immersion, that is, F_* is injective (its kernel is zero-dimensional):

$$
\begin{aligned}
F_*(\delta) = 0 \;\; &\Rightarrow \;\; \delta p^T + p\delta^T = 0 \\
&\Rightarrow \;\; p^T \left(\delta p^T + p\delta^T \right) = 0 \\
&\Rightarrow \;\; \delta^T = 0.
\end{aligned}
$$

We used the facts that $p^T \delta = 0$ and $p^T p = 1$.

Example 6 [A submersion onto the unit-sphere] Consider the map

$$
F \;:\; \mathbb{R}^n - \{0\} \to \mathsf{S}^{n-1}(\mathbb{R}) \qquad F(x) = \frac{x}{\|x\|}.
$$

We have the commutative diagram

It is easy to obtain the push-forward of $\widehat{F}$ at any point $p \in \mathbb{R}^n - \{0\}$:

$$
\widehat{F}_* \;:\; T_p \mathbb{R}^n - \{0\} \simeq \mathbb{R}^n \to T_{\widehat{F}(p)} \mathbb{R}^n \simeq \mathbb{R}^n
$$

$$
\widehat{F}_*(\delta) = \frac{1}{\|p\|} \left(I_n - \frac{pp^T}{\|p\|^2} \right) \delta = \frac{1}{\|p\|} \left(I_n - F(p)F(p)^T \right) \delta.
$$

On the other hand,

$$
T_{F(p)} \mathsf{S}^{n-1}(\mathbb{R}) \simeq \{ \gamma \in \mathbb{R}^n \;:\; F(p)^T \gamma = 0 \}.
$$

In conclusion, we can represent the push-forward $F_* \;:\; \mathbb{R}^n - \{0\} \to T_{F(p)} \mathsf{S}^{n-1}(\mathbb{R})$ by the linear map

$$
F_* \;:\; \mathbb{R}^n - \{0\} \to \{ \gamma \;:\; F(p)^T \gamma = 0 \} \qquad F_*(\delta) = \frac{1}{\|p\|} \left(I_n - F(p)F(p)^T \right) \delta.
$$

As an example, we can exploit the representation above to prove that the smooth map F is a submersion, that is, F_* is surjective: choose γ such that $F(p)^T \gamma = 0$. Letting $\delta = \|p\| \gamma$, we have $F_*(\delta) = \gamma$.

Example 7 [Another submersion onto the unit-sphere] Consider the map

$$F \ : \ \mathsf{O}(n) \to \mathsf{S}^{n-1}(\mathbb{R}) \qquad F(X) = Xe_1,$$

where $e_1 = (1, 0, \ldots, 0)^T$.

The map F is smooth and we have the commutative diagram where

$$
\begin{array}{ccc}
\mathsf{M}(n, \mathbb{R}) & \xrightarrow{\widehat{F}} & \mathbb{R}^n \\
\Big\uparrow{\scriptstyle \iota_{\mathsf{O}(n)}} & & \Big\uparrow{\scriptstyle \iota_{\mathsf{S}^{n-1}(\mathbb{R})}} \\
\mathsf{O}(n) & \xrightarrow{F} & \mathsf{S}^{n-1}(\mathbb{R})
\end{array}
$$

$\widehat{F}(X) = Xe_1$.

The push-forward of $\widehat{F}$ at any point $X_0 \in \mathsf{M}(n, \mathbb{R})$ is easily obtained:

$$\widehat{F}_* \ : \ T_{X_0}\mathsf{M}(n, \mathbb{R}) \simeq \mathsf{M}(n, \mathbb{R}) \to T_{\widehat{F}(X_0)}\mathsf{M}(n, \mathbb{R}) \simeq \mathsf{M}(n, \mathbb{R})$$

$$\widehat{F}_*(\Delta) = \Delta e_1.$$

Now, for $Q_0 = [\, q_1 \, q_2 \, \cdots \, q_n \,] \in \mathsf{O}(n)$,

$$
\begin{array}{rcl}
T_{Q_0}\mathsf{O}(n) & \simeq & Q_0 \mathsf{K}(n, \mathbb{R}) \\
T_{F(Q_0)}\mathsf{S}^{n-1}(\mathbb{R}) & \simeq & \{\delta \in \mathbb{R}^n \ \ q_1^T \delta = 0\}
\end{array}
$$

Thus, we can represent the push-forward $F_* \ : \ T_{Q_0}\mathsf{O}(n) \to T_{F(Q_0)}\mathsf{S}^{n-1}(\mathbb{R})$ by the linear map

$$F_* \ : \ Q_0 \mathsf{K}(n, \mathbb{R}) \to \{\delta \ : \ q_1^T \delta = 0\} \qquad F_*(Q_0 K) = Q_0 K e_1.$$

Exploiting the representation above, it is straightforward to show that F_* is surjective and, therefore, F is a submersion.

Lemma [Embedded submanifolds of product manifolds] If A is embedded in M and B is embedded in N, then $A \times B$ is embedded in $M \times N$.

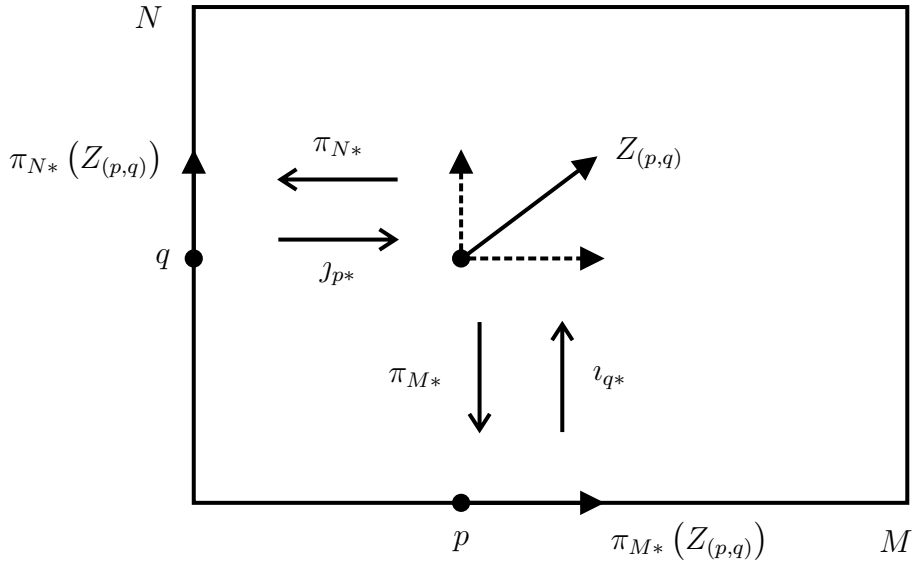**Example 1 [Tangent space identifications for product manifolds]**
Let $(p,q) \in M \times N$. We have the identification

$$T_{(p,q)} M \times N \simeq T_p M \oplus T_q N$$

due to the isomorphism

$$\pi_{M*} \times \pi_{N*} \: : \: T_{(p,q)} M \times N \to T_p M \oplus T_q N$$

$$\pi_{M*} \times \pi_{N*} \left(Z_{(p,q)} \right) = \left(\pi_{M*}(Z_{(p,q)}), \pi_{N*}(Z_{(p,q)}) \right).$$

The inverse map is given by

$$\imath_{q*} \oplus \jmath_{p*} \: : \: T_p M \oplus T_q N \to T_{(p,q)} M \times N \qquad \imath_{q*} \oplus \jmath_{p*}(X_p, Y_q) = \imath_{q*}(X_p) + \jmath_{p*}(Y_q),$$

where

$$
\begin{aligned}
\imath_q \: &: \: M \to M \times N \qquad x \mapsto (x, q) \\
\jmath_p \: &: \: N \to M \times N \qquad y \mapsto (p, y).
\end{aligned}
$$

The discussed identification $T_{(p,q)} M \times N \simeq T_p M \oplus T_q N$ can be used as follows. Suppose we have a smooth map $F : M \times N \to P$. We want to compute the push-forward of F at the point (p,q), that is, the linear map

$$F_* : T_{(p,q)} M \times N \to T_{F(p,q)} P \qquad Z_{(p,q)} \mapsto F_*(Z_{(p,q)}).$$

Since $T_{(p,q)} M \times N \simeq T_p M \oplus T_q N$, we know that it can be represented by a linear map

$$F_* : T_p M \oplus T_q N \to T_{F(p,q)} P \qquad (X_p, Y_q) \mapsto F_*(X_p, Y_q).$$

The next diagram illustrates the idea:

F_* (the identification)

To find out how to represent F_* by this latter map, we reason as follows:

$$
\begin{aligned}
F_*(X_p, Y_q) &\simeq F_* \circ \imath_{q*} \oplus \jmath_{p*}(X_p, Y_q) \\
&= F_*(\imath_{q*} X_p + \jmath_{p*} Y_q) \\
&= (F \circ \imath_q)_* X_p + (F \circ \jmath_p)_* Y_q \\
&= F_{q*} X_p + F_{p*} Y_q,
\end{aligned}
$$

where

$$F_q = F \circ \imath_q : M \to P \qquad x \mapsto F(x, q)$$

and

$$F_p = F \circ \jmath_p : M \to P \qquad y \mapsto F(p, y).$$

That is, F_q and F_p correspond to F when we hold fixed the 2nd and 1st argument at q and p, respectively.

For a specific example, let $M = N = P = \mathsf{M}(n, \mathbb{R})$ and consider the smooth map

$$F : M \times N \to P \qquad F(X, Y) = XY.$$

36

The push-forward of F at the point (X_0, Y_0) can be represented by the linear map

$$F_* : \mathsf{M}(n, \mathbb{R}) \oplus \mathsf{M}(n, \mathbb{R}) \to \mathsf{M}(n, \mathbb{R}) \qquad F_*(\Delta, \Psi) = \Delta Y_0 + X_0 \Psi.$$

Example 2 [Embedded submanifolds of product manifolds] Let A, B and C be embedded submanifolds of M, N and P, respectively. Let $F : A \times B \to C$ be a smooth map. Suppose that there exists a smooth map $\widehat{F} : M \times N \to P$ such that the following diagram commutes

$$
\begin{array}{ccc}
M \times N & \xrightarrow{\ \widehat{F}\ } & P \\[2pt]
\Big\uparrow{\iota_{A \times B}} & & \Big\uparrow{\iota_C} \\[2pt]
A \times B & \xrightarrow[\ F\ ]{} & C
\end{array}
$$

Note that $A \times B$ is embedded in $M \times N$. Thus, for any given $(a, b) \in A \times B$, we already know that we have the following diagram

$$
\begin{array}{ccc}
\iota_{A \times B *}\big(T_{(a,b)}A \times B\big) & \xrightarrow{\ \widehat{F}_*\ } & \iota_{C *}\big(T_{F(a,b)}C\big) \\[2pt]
\Big\uparrow{\iota_{A \times B *}} & & \Big\uparrow{\iota_{C *}} \\[2pt]
T_{(a,b)}A \times B & \xrightarrow[\ F_*\ ]{} & T_{F(a,b)}C
\end{array}
$$

which allows us to represent the "hard" linear map

$$F_* : T_{(a,b)}A \times B \to C$$

by the "easier" one

$$\widehat{F}_* : \iota_{A \times B *}\big(T_{(a,b)}A \times B\big) \subset T_{(a,b)}M \times N \to \iota_{C *}\big(T_{F(a,b)}C\big) \subset T_{F(a,b)}P.$$

Our goal here is to exploit the tangent space identifications discussed in the previous example to find out another representation for F_*.

We start by noting that we have the following two diagrams

$$
\begin{array}{ccc}
M & \xrightarrow{\;\widehat{\imath}_b\;} & M \times N \\
\Big\uparrow{\scriptstyle \iota_A} & & \Big\uparrow{\scriptstyle \iota_{A\times B}} \\
A & \xrightarrow{\;\imath_b\;} & A \times B
\end{array}
\qquad
\begin{array}{ccc}
N & \xrightarrow{\;\widehat{\jmath}_a\;} & M \times N \\
\Big\uparrow{\scriptstyle \iota_B} & & \Big\uparrow{\scriptstyle \iota_{A\times B}} \\
B & \xrightarrow{\;\jmath_a\;} & A \times B
\end{array}
$$

where

$$
\begin{aligned}
\imath_b &: A \to A \times B & x &\mapsto (x,b) \\
\widehat{\imath}_b &: M \to M \times N & x &\mapsto (x,b) \\
\jmath_a &: B \to A \times B & y &\mapsto (a,y) \\
\widehat{\jmath}_a &: N \to M \times N & y &\mapsto (a,y).
\end{aligned}
$$

From this, it follows that

$$
\begin{array}{ccc}
\iota_{A*}\left(T_a A\right) \oplus \iota_{B*}\left(T_b B\right) & \xrightarrow{\;\widehat{\imath}_{b*} \oplus \widehat{\jmath}_{a*}\;} & \iota_{A\times B*}\left(T_{(a,b)} A \times B\right) \\
\Big\uparrow{\scriptstyle \iota_{A*} \times \iota_{B*}} & & \Big\uparrow{\scriptstyle \iota_{A\times B*}} \\
T_a A \oplus T_b B & \xrightarrow{\;\imath_{b*} \oplus \jmath_{a*}\;} & T_{(a,b)} A \times B
\end{array}
$$

where

$$
\begin{aligned}
\imath_{b*} \oplus \jmath_{a*} &: T_a A \oplus T_b B \to T_{(a,b)} A \times B & (X_a, Y_b) &\mapsto \imath_{b*}(X_a) + \jmath_{a*}(Y_b) \\
\widehat{\imath}_{b*} \oplus \widehat{\jmath}_{a*} &: T_a M \oplus T_b N \to T_{(a,b)} M \times N & (X_a, Y_b) &\mapsto \widehat{\imath}_{b*}(X_a) + \widehat{\jmath}_{a*}(Y_b) \\
\iota_{A*} \times \iota_{B*} &: T_a A \oplus T_b B \to T_a M \oplus T_b N & (X_a, Y_b) &\mapsto (\iota_{A*}(X_a), \iota_{B*}(Y_b))
\end{aligned}
$$

The equality

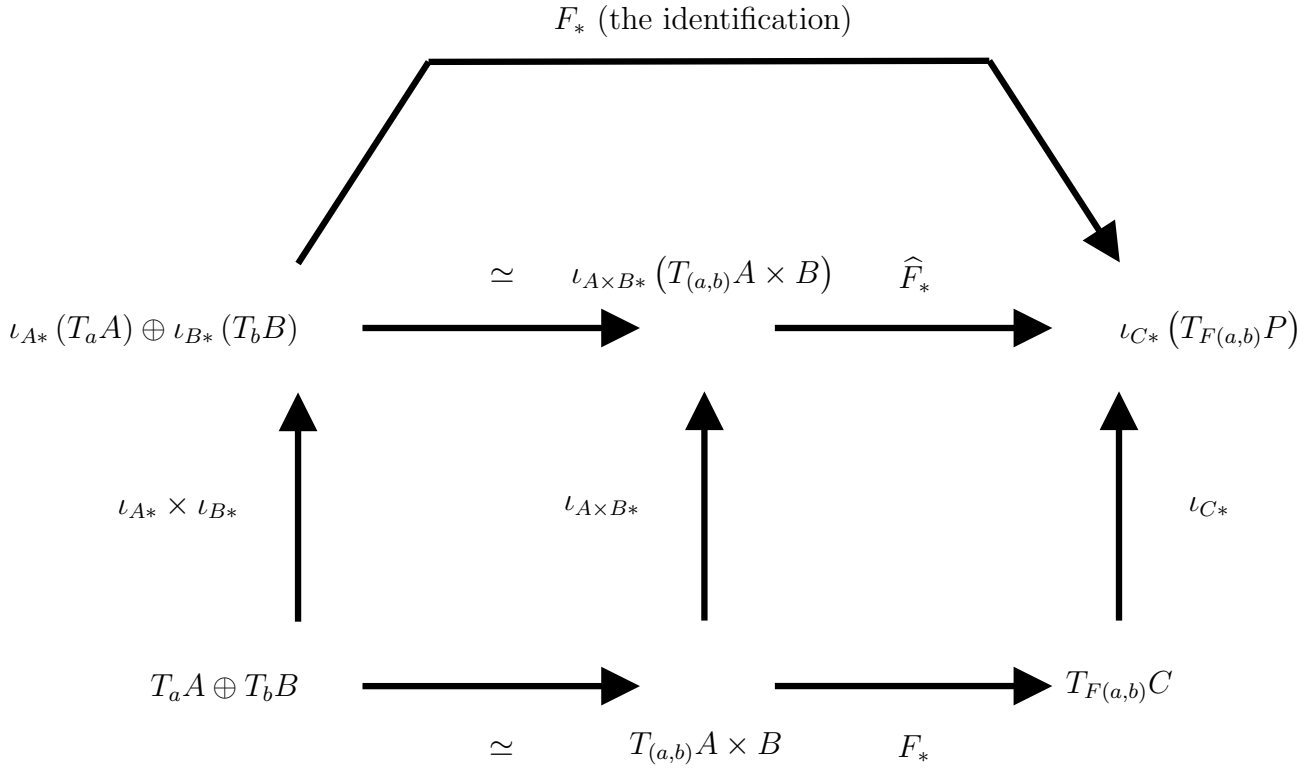$$\iota_{A\times B*} \circ \imath_{b*} \oplus \jmath_{a*} = \widehat{\imath}_{b*} \oplus \widehat{\jmath}_{a*} \circ \iota_{A*} \times \iota_{B*}$$

expressed in the last diagram can be proved as follows:

$$
\begin{aligned}
\iota_{A\times B*} \circ \imath_{b*} \oplus \jmath_{a*}(X_a, Y_b) &= \iota_{A\times B*}\left(\imath_{b*}(X_a) + \jmath_{a*}(Y_b)\right) \\
&= \iota_{A\times B*} \circ \imath_{b*}(X_a) + \iota_{A\times B*} \circ \jmath_{a*}(Y_b) \\
&\overset{(a)}{=} \left(\iota_{A\times B} \circ \imath_b\right)_* (X_a) + \left(\iota_{A\times B} \circ \jmath_a\right)_* (Y_b) \\
&\overset{(b)}{=} \left(\widehat{\imath}_b \circ \iota_A\right)_* (X_a) + \left(\widehat{\jmath}_a \circ \iota_B\right)_* (Y_b) \\
&= \widehat{\imath}_{b*}\left(\iota_{A*}(X_a)\right) + \widehat{\jmath}_{a*}\left(\iota_{B*}(Y_b)\right) \\
&= \widehat{\imath}_{b*} \oplus \widehat{\jmath}_{a*}\left(\iota_{A*}(X_a), \iota_{B*}(Y_b)\right) \\
&= \widehat{\imath}_{b*} \oplus \widehat{\jmath}_{a*} \circ \iota_{A*} \times \iota_{B*}(X_a, Y_b).
\end{aligned}
$$

In (a), the chain rule for push-forwards was used. In (b), we used the two commutative diagrams in page 38.

Now, taking the last diagram in page 38 and plugging it on the left of the last diagram in page 37 yields

$$F_* \text{ (the identification)}$$

$$\iota_{A*}\left(T_a A\right) \oplus \iota_{B*}\left(T_b B\right) \quad \overset{\simeq}{\longrightarrow} \quad \iota_{A\times B*}\left(T_{(a,b)} A \times B\right) \quad \overset{\widehat{F}_*}{\longrightarrow} \quad \iota_{C*}\left(T_{F(a,b)} P\right)$$

$$\iota_{A*} \times \iota_{B*} \uparrow \qquad \iota_{A\times B*} \uparrow \qquad \iota_{C*} \uparrow$$

$$T_a A \oplus T_b B \longrightarrow T_{(a,b)} A \times B \longrightarrow T_{F(a,b)} C$$

$$\simeq \qquad F_*$$

Since the arrows marked with $\simeq$ denote isomorphisms, this shows that F_* can be represented by the linear map

$$F_* \,:\, \iota_{A*}\left(T_a A\right) \oplus \iota_{B*}\left(T_b B\right) \to \iota_{C*}\left(T_{F(a,b)} C\right)$$

given by

$$\begin{aligned}
\left(\iota_{A*}(X_a), \iota_{B*}(Y_b)\right) \;\mapsto\; & \widehat{F}_*\left(\widehat{\imath}_{b*}\left(\iota_{A*}(X_a)\right) + \widehat{\jmath}_{a*}\left(\iota_{B*}(Y_b)\right)\right) \\
=\; & \widehat{F}_{b*}\left(\iota_{A*}(X_a)\right) + \widehat{F}_{a*}\left(\iota_{B*}(Y_b)\right),
\end{aligned}$$

where

$$\begin{aligned}
\widehat{F}_b \,:\, M \to P \qquad & x \mapsto \widehat{F}(x, b) \\
\widehat{F}_a \,:\, N \to P \qquad & y \mapsto \widehat{F}(a, y).
\end{aligned}$$

Example 3 [Polar decomposition is a diffeomorphism] Consider the map

$$F \,:\, \mathsf{P}(n, \mathbb{R}) \times \mathsf{O}(n) \to \mathsf{GL}(n, \mathbb{R}) \qquad F(P, Q) = PQ.$$

The map F is smooth (why?).

We have the following commutative diagram

$$
\begin{array}{ccc}
\mathsf{M}(n,\mathbb{R}) \times \mathsf{M}(n,\mathbb{R}) & \xrightarrow{\ \widehat{F}\ } & \mathsf{M}(n,\mathbb{R}) \\[2mm]
\Big\uparrow{\scriptstyle \iota_{\mathsf{P}(n,\mathbb{R})\times\mathsf{O}(n)}} & & \Big\uparrow{\scriptstyle \iota_{\mathsf{GL}(n,\mathbb{R})}} \\[2mm]
\mathsf{P}(n,\mathbb{R}) \times \mathsf{O}(n) & \xrightarrow[\ F\ ]{} & \mathsf{GL}(n,\mathbb{R})
\end{array}
$$

where

$$
\widehat{F} \ : \ \mathsf{M}(n,\mathbb{R}) \times \mathsf{M}(n,\mathbb{R}) \to \mathsf{M}(n,\mathbb{R}) \qquad \widehat{F}(X,Y) = XY.
$$

Note that $\mathsf{P}(n,\mathbb{R})$, $\mathsf{O}(n)$ and $\mathsf{GL}(n,\mathbb{R})$ are embedded submanifolds of $\mathsf{M}(n,\mathbb{R})$.

Also:
$$
\begin{array}{rcl}
\iota_{\mathsf{P}(n,\mathbb{R})*}\left(T_{P_0}\mathsf{P}(n,\mathbb{R})\right) & \simeq & \mathsf{S}(n,\mathbb{R}) \\
\iota_{\mathsf{O}(n)*}\left(T_{Q_0}\mathsf{O}(n)\right) & \simeq & Q_0\mathsf{K}(n,\mathbb{R}) \\
\iota_{\mathsf{GL}(n,\mathbb{R}))*}\left(T_{X_0}\mathsf{GL}(n,\mathbb{R})\right) & \simeq & \mathsf{M}(n,\mathbb{R}).
\end{array}
$$

Thus, the push-forward

$$
F_* \ : \ T_{(P_0,Q_0)}\mathsf{P}(n,\mathbb{R}) \times \mathsf{O}(n) \to T_{F(P_0,Q_0)}\mathsf{GL}(n,\mathbb{R})
$$

can be represented by the linear map

$$
F_* \ : \ \mathsf{S}(n,\mathbb{R}) \oplus Q_0\mathsf{K}(n,\mathbb{R}) \to \mathsf{M}(n,\mathbb{R}) \qquad F_*(\Delta, \Psi) = \Delta Q_0 + P_0\Psi.
$$

Since F_* is an injective linear map for any (P_0, Q_0), we conclude that F is a diffeomorphism.

References

[1] J. Lee, *Introduction to Smooth Manifolds*, Springer-Verlag, 2000.