CHAPTER 12

Comparison of Two Factorization Methods: Cross Section Versus Proper Action

The factorization method of Andersson, Brøns, and Jensen (1983) was mentioned briefly in Chapter 1. For brevity we shall term this the "ABJ" method, whereas the method that uses construction of a cross section—as developed in this monograph—will be designated by "W." It was mentioned in Chapter 1 that the main assumption made in the ABJ method is that the l.c. group G acts *properly* on the l.c. space $\mathfrak{X}$ (Definition 2.3.6). In the W method, properness of the action of G, although not explicitly assumed, is implied by Assumption 8.2. For, $\mathfrak{X}$ is homeomorphic to $\mathcal{Y} \times \mathfrak{T}$ by Proposition 8.4, and G acts properly on $\mathcal{Y} \times \mathfrak{T}$ since it acts trivially on $\mathfrak{T}$ and properly on $\mathcal{Y}$ by the compactness of G_0 (Proposition 2.3.11). Thus, both the ABJ and the W method rely on properness of the action of G. Below we shall concentrate on differences between the two methods.

In the ABJ method it is further assumed that there is another l.c. space $\mathcal{Y}$ on which G acts transitively and properly, and that there is a continuous and equivariant function $u : \mathfrak{X} \to \mathcal{Y}$, where u represents some statistic of interest. (Actually, this function is denoted t in ABJ. We have changed the notation from t to u in order to avoid confusion with the maximal invariant t of Chapter 8.) For instance, in the problem treated in Section 10.7 (Section 2 in ABJ) $\mathcal{Y}$ is the space of all $PD(2p)$ matrices of complex structure, and the equivariant function $u : \mathcal{X} \to \mathcal{Y}$ is defined by

$$u(S) = \frac{1}{2}\begin{bmatrix} S_{11} + S_{22} & S_{12} - S_{21} \\ S_{21} - S_{12} & S_{11} + S_{22} \end{bmatrix}. \tag{12.1}$$

This is a constant times the maximum likelihood estimator of Σ if the hypothesis is true. Let there be given on $\mathcal{X}$ a relatively invariant measure λ with multiplier χ. (In ABJ λ is chosen invariant, but the extension to the slightly more general case of a relatively invariant measure is no harder and makes the comparison with the W method easier.) Choose any $y_0 \in \mathcal{Y}$ and let G_1 be the isotropy subgroup of G at y_0. Then $\mathcal{Y}$ and G/G_1 are homeomorphic by Corollary 2.3.15 (change X and x there to $\mathcal{Y}$ and y_0 here). Thus, any measure on G/G_1 may be transferred to $\mathcal{Y}$. Let μ_1 be a fixed χ-relatively invariant measure on G/G_1, therefore on $\mathcal{Y}$, the existence of which is guaranteed by Corollary 7.4.4. Then an arbitrary χ-relative invariant measure on $\mathcal{Y}$ is of the form $c\mu_1$, with $c > 0$, again by Corollary 7.4.4. Let $\pi : \mathcal{X} \to \mathcal{X}/G$ be the orbit projection. Note that $\mathcal{X}/G$ is l.c., by Theorem 2.3.13(a). G acts on $\mathcal{Y} \times \mathcal{X}/G$ by $g(y, z) = (gy, z)$ for $y \in \mathcal{Y}$, $z \in \mathcal{X}/G$. The function $(u, \pi) : \mathcal{X} \to \mathcal{Y} \times \mathcal{X}/G$ is equivariant since $(u, \pi)(gx) = (u(gx), \pi(gx)) = (gu(x), \pi(x)) = g(u(x), \pi(x))$, using the equivariance of u and the definition of the action of G on $\mathcal{Y} \times \mathcal{X}/G$. It is shown by Andersson, Brøns, and Jensen (1983), in their Lemma 3, that (u, π) is a proper mapping and it follows that the image of λ under (u, π), say $\mu = (u, \pi)(\lambda)$, is a measure on $\mathcal{Y} \times \mathcal{X}/G$ (see Section 6.3, in particular (6.3.4)). The following computation shows that μ is χ-relatively invariant: Let $f \in \mathcal{K}(\mathcal{Y} \times \mathcal{X}/G)$ and put $f^* = f \circ (u, \pi)$, then $f^* \in \mathcal{K}(\mathcal{X})$ and $\int f d\mu = \int f^* d\lambda$ by definition of the induced measure μ. A simple computation shows $gf^* = gf \circ (u, \pi)$, for $g \in G$. Then $\int gf d\mu = \int gf^* d\lambda = \chi(g) \int f^* d\lambda$ (since λ is χ-relatively invariant) $= \chi(g) \int f d\mu$. By the χ-relative invariance of μ and Theorem 7.5.1 there is a measure μ_2 on $\mathcal{X}/G$ such that $\mu = \mu_1 \otimes \mu_2$. This is essentially the conclusion of Lemma 3 in ABJ (their ν_0 and κ are μ_1 and μ_2 here). In their applications the random variable X has a distribution of the form $f_1(u(x))f_2(\pi(x))\lambda(dx)$. Then the factorization of $(u,\pi)(\lambda)$ shows that $u(X)$ and $\pi(X)$ are independent, with distributions $f_1(y)\mu_1(dy)$ and $f_2(z)\mu_2(dz)$, respectively,. However, an explicit expression for μ_2 is not found by this method but has to be obtained by other means.

We shall now compare the ABJ and W methods. In the latter the space $\mathcal{Y}$ is always taken to be G/G_0, where the compact group G_0 is the common value of all G_z when z traverses the cross section $\mathcal{Z}$, and the function u of ABJ is the function y of (8.3). For notational clarity we shall denote this function by a different symbol:

$$\eta(x) = [g] \quad \text{if} \quad x = gz, \quad z \in \mathcal{Z}, \tag{12.2}$$

where $[g]$ stands for $gG_0 \in \mathcal{Y}$. If $\mathcal{Z}$ is taken as a representation of $\mathcal{X}/G$, then the orbit projection is

$$\pi(x) = z \quad \text{if} \quad x = gz, \quad z \in \mathcal{Z}. \tag{12.3}$$

If the space $\mathcal{T}$ of Chapter 8 is identified with $\mathcal{Z}$ then the diffeomorphism φ of (8.7) may be considered a function $\mathcal{Y} \times \mathcal{Z} \to \mathcal{X}$, and its inverse is

$$\varphi^{-1} = (\eta, \pi). \tag{12.4}$$

Thus, in the W method there is a space $\mathcal{Y}$ (namely, G/G_0) on which G acts transitively and properly (Proposition 2.3.11), and a continuous and equivariant function $\eta : \mathcal{X} \to \mathcal{Y}$. Then the factorization of $(\eta,\pi)(\lambda)$ by the W method, essentially given by (8.10), is a special case of the factorization of $(u,\pi)(\lambda)$ by the ABJ method. It will be shown now that, conversely, the ABJ factorization can be derived from the W factorization provided the following condition is satisfied: u maps $\mathcal{Z}$ into a single point of $\mathcal{Y}$ (i.e., u is constant on $\mathcal{Z}$). In the theorem below we shall not need all the regularity of Assumption 8.2, but we shall return to that regularity when applying the theorem to the comparison of the ABJ and W factorizations.

12.1. THEOREM. *Let $\mathcal{X}$ be a l.c. space and G a l.c. group acting continuously on $\mathcal{X}$. Suppose there is a global cross section $\mathcal{Z} \subset \mathcal{X}$ for* *which there exists a closed subgroup G_0 of G such that $G_z = G_0$ for every $z \in \mathcal{Z}$. Define $\eta : \mathcal{X} \to G/G_0$ by (12.2), so that η is equivariant. Suppose there is a function u that maps $\mathcal{X}$ into some other l.c. space $\mathcal{Y}$ on which G acts continuously and transitively, such that u is equivariant and $u(\mathcal{Z}) = y_0$ for some $y_0 \in \mathcal{Y}$. Denote the isotropy subgroup G_{y_0} by G_1 and assume that $\mathcal{Y}$ and G/G_1 are homeomorphic. Then there exists a continuous, open, and equivariant function $\psi : G/G_0 \to \mathcal{Y}$ such that*

$$(12.5) \qquad u = \psi \circ \eta.$$

In fact, if $\mathcal{Y}$ and G/G_1 are identified, then ψ is the function $G/G_0 \to G/G_1$ defined by

$$(12.6) \qquad \psi(gG_0) = gG_1, \quad g \in G.$$

If the stronger assumption is made that G act properly on $\mathcal{Y}$ (from which the homeomorphism of $\mathcal{Y}$ and G/G_1 is a consequence, by Corollary 2.3.15) then ψ is proper.

PROOF. First we verify that $G_0 \subset G_1$: take any $z \in \mathcal{Z}$, then $g \in G_0$ implies $y_0 = u(z) = u(gz) = gu(z) = gy_0$ so that $g \in G_1$. It follows that ψ of (12.6) is well defined. Consider, for $i = 0, 1$, the coset projection $\pi_i : G \to G/G_i$, which is continuous, open, and onto. Then we have $\pi_1 = \psi \circ \pi_0$. If A is an arbitrary subset of G/G_1, then $\psi^{-1}(A) = \pi_0(\pi_1^{-1}A)$ since π_0 is onto (cf. (2.1.9)). If A is open, then the continuity of π_1 and openness of π_0 guarantee that $\psi^{-1}(A)$ is open. Hence, ψ is continuous. If B is an arbitrary subset of G/G_0, then $\psi(B) = \pi_1(\pi_0^{-1}B)$, again because π_0 is onto (cf. (2.1.10)). If B is open, then the continuity of π_0 and openness of π_1 guarantee that $\psi(B)$ is open. Hence ψ is an open map. The equivariance of ψ is immediate by its definition (12.6).

Next, we show the validity of (12.5). It is sufficient to check this separately on each G-orbit in $\mathcal{X}$. Take $z \in \mathcal{Z}$ arbitrarily and consider the orbit Gz in $\mathcal{X}$. Then G is transitive both over Gz and over $\mathcal{Y}$, so that an equivariant function $Gz \to \mathcal{Y}$ is determined by its value at one point. Now u is given to be equivariant and $u(z) = y_0$. Also, both ψ and η are equivariant, so that $\psi \circ \eta$ is equivariant. Furthermore, by taking $g = e$ in (12.2) and (12.6), we see that $(\psi \circ \eta)(z) = \psi(G_0) = G_1 \in G/G_1$, which corresponds to $y_0 \in \mathcal{Y}$. Thus, under the identification of $\mathcal{Y}$ and G/G_1, the equivariant functions u and $\psi \circ \eta$ are equal at one point of Gz and are therefore equal along the whole orbit.

Finally, if G acts properly on $\mathcal{Y} = Gy_0$, then G_1 is compact (Theorem 2.3.13(c)) and $\mathcal{Y}$ and G/G_1 are homeomorphic (Corollary 2.3.15). Since $G_0 \subset G_1$, G_0 is also compact. Furthermore, by the compactness of G_i, π_i is proper, $i = 0, 1$ (Proposition 2.3.5). For any compact $A \subset G/G_1$, $\pi_1^{-1}(A)$ is a compact subset of G, by Theorem 2.2.3. Then $\psi^{-1}(A) = \pi_0(\pi_1^{-1}A)$ is a compact subset of G/G_0 which shows that ψ is proper, again by Theorem 2.2.3. □

12.2. REMARK. Define $\mathcal{E}$ to be the set of all pairs $(u, \mathcal{Y})$ satisfying the hypotheses of Theorem 12.1 (not including the assumption of proper action of G on $\mathcal{Y}$). Then $(\eta, G/G_0) \in \mathcal{E}$ (with $G_1 = G_0$ and $\psi =$ identity map). For an arbitrary $(u, \mathcal{Y}) \in \mathcal{E}$ the conclusion (12.5) can be phrased: u depends on x through $\eta(x)$. Thus, we are justified in describing $(\eta, G/G_0)$ as a **maximal equivariant** function in the set $\mathcal{E}$. □

12.3. REMARK. In Theorem 12.1 it is not assumed that η is continuous. If η is continuous (e.g., if Assumption 8.2 is satisfied), then every u with $(u, \mathcal{Y}) \in \mathcal{E}$ must be continuous, by (12.5). However, if it is only given that for some $(u, \mathcal{Y}) \in \mathcal{E}$, u is continuous, then it does not necessarily follow that η is continuous. (For instance, $\mathcal{Y}$ could be a single point.) □

12.4. REMARK. In Theorem 12.1 it is not assumed either that G acts properly on $\mathcal{X}$. But if $(u, \mathcal{Y}) \in \mathcal{E}$ with u continuous, and if G acts properly on $\mathcal{Y}$, then G must act properly on $\mathcal{X}$. This follows from $((A, B)) \subset ((u(A), u(B)))$ (for notation see (2.3.2)) if A and B are arbitrary subsets of $\mathcal{X}$, using the equivariance of u. Then take A, B compact, use the continuity of u, and the conclusion follows from Proposition 2.3.8. □

We return now to the comparison of the ABJ and W factorizations. Suppose Assumption 8.2 is satisfied so that there is a cross section $\mathcal{Z}$ and compact $G_0 \subset G$ with $G_z = G_0$ for every $z \in \mathcal{Z}$. Theorem 8.6 leads to a factorization (8.10), which is of the form $\varphi^{-1}(\lambda) = \mu_0 \otimes \mu_2$, where μ_0 is a measure on G/G_0. Suppose that the continuous and equivariant function u of ABJ (actually, their t) maps $\mathcal{Z}$ into a single point of $u(\mathcal{X}) = \mathcal{Y}$. Then Theorem 12.1 applies. Since it is also assumed by ABJ that G acts properly on $\mathcal{Y}$, the conclusion is (12.5) with ψ proper. It follows that the ABJ factorization is

$$(u, \pi)(\lambda) = \psi(\mu_0) \otimes \mu_2. \tag{12.7}$$

12.5. EXAMPLES. In the problem of Section 10.7 (the first problem in ABJ, Section 2) we had a cross section $\mathcal{Z}$ consisting of all matrices $\operatorname{diag}(I_p + L, I_p - L)$, $L = \operatorname{diag}(\ell_1, \dots, \ell_p)$, with $1 > \ell_1 > \cdots > \ell_p > 0$. The function u is defined in (12.1) and it is seen that for any $z \in \mathcal{Z}$, $u(z) = \operatorname{diag}(I_p, I_p)$. Thus, u is constant on $\mathcal{Z}$ so that Theorem 12.1 applies.

In the second problem of Andersson, Brøns, and Jensen (1983), Section 3 (not treated in Chapter 10, but treated in Wijsman, 1986, Section 7.7(b)) $\mathcal{X}$ consists of $2p \times 2p$ positive definite matrices $S = ((S_{ij}))$, $i, j = 1, 2$, with $S_{11} = S_{22}$, $S_{21} = S_{12}$, and all $S_{ij} : p \times p$. The function u can be described by stating that $u(S)$ equals S with S_{12} and S_{21} set equal to 0. A cross section $\mathcal{Z}$ may be taken as the range of the function s that maps $(\omega_1, \dots, \omega_r)$ (with $1 > \omega_1 > \cdots > \omega_r > 0$ and $r = [p/2]$) into a matrix S with $S_{11} = S_{22} = I_p$ and $S_{21} = S_{12} = \Omega$, where Ω depends on the ω_i in a manner that is unimportant for the present considerations. Thus, it is seen that if $S \in \mathcal{Z}$, then $u(S) = \operatorname{diag}(I_p, I_p) =$ constant so that again Theorem 12.1 applies. □

In summary, under an appropriate amount of regularity, the ABJ and W factorization results are obtainable from each other. That is, starting from a relatively invariant measure λ on $\mathcal{X}$, the ABJ factorization $(u, \pi)(\lambda) = \mu_1 \otimes \mu_2$ on $\mathcal{Y} \times \mathcal{Z}$ and the W factorization

$(\eta, \pi)(\lambda) = \mu_0 \otimes \mu_2$ on $G/G_0 \times \mathcal{Z}$ are equivalent. A major difference between the two methods is that the ABJ method requires fewer assumptions, but the W method has a built-in capability of obtaining an explicit expression for the measure μ_2 on $\mathcal{Z}$ (or $\mathcal{T}$ of Chapter 8).