Supplemental file for the paper titled "A Multivariate Sign EWMA Control Chart"

supplement.pdf This pdf file provides certain technical details, including proofs of Propositions 1-4 in Section 2, the algorithm for computing $\widehat{\boldsymbol{\theta}}_0$ and $\widehat{\mathbf{A}}_0$, and the AEC dataset used in the paper.

Proofs of Propositions in Section 2.2

Proof of Proposition 1. To prove that the affine invariance of MSEWMA is equivalent to showing that for any $p \times p$ nonsingular matrix $\mathbf{D}$, the charting statistics, Q_i, based on $\boldsymbol{x}_i$ and $\boldsymbol{y}_i = \mathbf{D}\boldsymbol{x}_i$ are the same. In what follows, we use the superscripts "$(\boldsymbol{x})$" and "$(\boldsymbol{y})$" to distinguish the corresponding statistics or parameters based on the sample $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$, i.e., for example $Q_i^{(\boldsymbol{x})}$ and $Q_i^{(\boldsymbol{y})}$.

First of all, according to a Proposition in Hettmansperger and Randles (2002), we know the AEM-median, $\boldsymbol{\theta}_0$, is affine equivariant, say $\boldsymbol{\theta}_0^{(\boldsymbol{y})} = \mathbf{D}\boldsymbol{\theta}_0^{(\boldsymbol{x})}$. Furthermore, by the similar arguments in the Appendix of Randles (2000), which shows the affine invariance of Q in Eq.(6), we can easily show that

$$\mathbf{A}_0^{(y)} = c\mathbf{A}_0^{(x)}\mathbf{D}^{-1},$$

where c is a positive constant to normalize $\mathbf{A}_0^{(y)}$ so that its upper left-hand element equals one. Then, from the definition of $\boldsymbol{v}_i$ in Eq.(11), we have

$$\boldsymbol{v}_i^{(\boldsymbol{y})} = \frac{\mathbf{A}_0^{(\boldsymbol{y})}(\boldsymbol{y}_i - \boldsymbol{\theta}_0^{(\boldsymbol{y})})}{||\mathbf{A}_0^{(\boldsymbol{y})}(\boldsymbol{y}_i - \boldsymbol{\theta}_0^{(\boldsymbol{y})})||} = \frac{\mathbf{A}_0^{(\boldsymbol{x})}\mathbf{D}^{-1}(\mathbf{D}\boldsymbol{x}_i - \mathbf{D}\boldsymbol{\theta}_0^{(\boldsymbol{x})})}{||\mathbf{A}_0^{(\boldsymbol{x})}\mathbf{D}^{-1}(\mathbf{D}\boldsymbol{x}_i - \mathbf{D}\boldsymbol{\theta}_0^{(\boldsymbol{x})})||} = \boldsymbol{v}_i^{(\boldsymbol{x})}.$$

Now, the proposition immediately follows from the definitions of w_i and Q_i. $\square$

Proof of Proposition 2. Recall the definition of distributions with elliptical directions introduced in Section 2.1. Based on Proposition 1, Q_i is not influenced by arbitrary $p \times p$ nonsingular $\mathbf{D}$. Note that Q_i uses only the direction of an observation from the origin and does not use its distance from the origin. That is, if $\boldsymbol{x}_i = r_i \boldsymbol{u}_i$, where $u_i = \boldsymbol{x}_i / ||\boldsymbol{x}_i||$, then it is apparent that r_i plays no role in $\boldsymbol{v}_i$ of Eq.(11) and thus it does not affect the value of Q_i. Therefore, the proposition follows from similar arguments in the proof of Proposition 1 and the details are omitted here for the sake of brevity. $\square$

Proof of Proposition 3. Based on Proposition 2, without loss of generality, we assume that $\boldsymbol{x}_i$ are i.i.d. standard p-dimensional multinormal variables. Then $\boldsymbol{v}_i = \mathbf{A}_0(\boldsymbol{x}_i - \boldsymbol{\theta}_0)/||\mathbf{A}_0(\boldsymbol{x}_i - \boldsymbol{\theta}_0)|| = \boldsymbol{x}_i/||\boldsymbol{x}_i||$, where we use $\boldsymbol{\theta}_0 = \mathbf{0}$ and $\mathbf{A}_0 = \mathbf{I}_p$. $\boldsymbol{v}_i$ is a random variable with a spherical (uniform) distribution on $S(1)$, where $S(r)$ denotes the p-dimensional sphere of radius $r > 0$. Denote $a = (2 - \lambda)p/\lambda$. Note that

$$\Pr\left\{Q_i < x | Q_1, \ldots, Q_{i-1}\right\} = \Pr\left\{a||\lambda \boldsymbol{v}_i + (1-\lambda)\boldsymbol{w}_{i-1}||^2 < x | Q_1, \ldots, Q_{i-1}\right\}.$$

Similar to the proof of Proposition 1 in Runger and Prabhu (1996), by induction ,we can show that the distribution of $\boldsymbol{w}_{i-1}$ given $||\boldsymbol{w}_1||, ||\boldsymbol{w}_2||, \ldots, ||\boldsymbol{w}_{i-1}||$ is uniform on $S(||\boldsymbol{w}_{i-1}||)$. It follows that

$$\Pr\left\{Q_i < x | Q_1, \ldots, Q_{i-1}\right\} = \Pr\left\{a||\lambda \boldsymbol{v}_i + (1-\lambda)||\boldsymbol{w}_{i-1}||\boldsymbol{u}||^2 < x | Q_{i-1}\right\},$$

where $\boldsymbol{u}$ denotes the spherical variable on $S(1)$. This completes the proof. $\square$

Proof of Proposition 4. Note that $E(\boldsymbol{v}_i) = 0$ and $\mathrm{Cov}(\boldsymbol{v}_i) = p^{-1}\mathbf{I}_p$. Rewrite $\boldsymbol{w}_i$ as $\boldsymbol{w}_i = \lambda \sum_{j=1}^{i} (1-\lambda)^{i-j} \boldsymbol{v}_j$. By using the fact that the $\boldsymbol{v}_j$'s are i.i.d., this proposition follows immediately from the Lindeberg-Feller central limit theorem. $\square$

Computing $\widehat{\boldsymbol{\theta}}_0$ and $\widehat{\mathbf{A}}_0$

The iterative algorithm given by Hettmansperger and Randles (2002) for computing $(\widehat{\boldsymbol{\theta}}_0, \widehat{\mathbf{A}}_0)$ from the IC reference dataset, $\{\boldsymbol{x}_{-m_0+1}, \ldots, \boldsymbol{x}_0\}$, is described as follows:

1. Start by finding initial values for $\widehat{\boldsymbol{\theta}}_0$, denoted as $\widehat{\boldsymbol{\theta}}_0^{(0)}$. We recommend using the multivariate L_1 median, that is

$$\widehat{\boldsymbol{\theta}}_0^{(0)} = \arg\min_{\boldsymbol{\theta}} \sum_{i=-m_0+1}^{0} ||\boldsymbol{x}_i - \boldsymbol{\theta}||.$$

This minimization problem can be easily implemented by the modified simplex algorithm illustrated by Bedall and Zimmermann (1979).

2. At the lth iteration, for $l \geq 0$, find the value for $\widehat{\mathbf{A}}_0^{(l)}$ given $\widehat{\boldsymbol{\theta}}_0^{(l)}$, which satisfies

$$\frac{1}{m_0} \sum_{i=-m_0+1}^{0} \left(\frac{\mathbf{A}_0^{(l)}(\boldsymbol{x}_i - \widehat{\boldsymbol{\theta}}_0^{(l)})(\boldsymbol{x}_i - \widehat{\boldsymbol{\theta}}_0^{(l)})'\mathbf{A}_0'^{(l)})}{||\mathbf{A}_0^{(l)}(\boldsymbol{x}_i - \widehat{\boldsymbol{\theta}}_0^{(l)})||^2} \right) = \frac{1}{p}\mathbf{I}_p.$$

The algorithm for finding $\widehat{\mathbf{A}}_0^{(l)}$ also involves an iterative procedure described by Tyler (1987):

   (a) Starting with the initial value $\boldsymbol{\Omega} = \mathbf{I}_p$;

   (b) Set $\boldsymbol{\Omega}_x = [p/\text{trace}(\boldsymbol{\Omega})]\boldsymbol{\Omega}$;

   (c) Choose $\mathbf{A}_\Omega$ so that $\mathbf{A}'_\Omega \mathbf{A}_\Omega = \boldsymbol{\Omega}_x^{-1}$. This can be done by finding the upper triangular Choleski factorization of $\boldsymbol{\Omega}_x^{-1}$, divided by the upper left-hand element of that upper triangular matrix.

   (d) Use an iteration step that transforms from one $\boldsymbol{\Omega}$ to the next by

$$\boldsymbol{\Omega} \leftarrow p\boldsymbol{\Omega}^{\frac{1}{2}} \frac{1}{m_0} \sum_{i=-m_0+1}^{0} \left(\frac{\mathbf{A}_\Omega(\boldsymbol{x}_i - \widehat{\boldsymbol{\theta}}_0^{(l)})}{||\mathbf{A}_\Omega(\boldsymbol{x}_i - \widehat{\boldsymbol{\theta}}_0^{(l)})||} \right) \left(\frac{\mathbf{A}_\Omega(\boldsymbol{x}_i - \widehat{\boldsymbol{\theta}}_0^{(l)})}{||\mathbf{A}_\Omega(\boldsymbol{x}_i - \widehat{\boldsymbol{\theta}}_0^{(l)})||} \right)' \boldsymbol{\Omega}^{\frac{1}{2}},$$

   (e) Repeat Steps b-d until convergence.

3. Update the estimation of $\boldsymbol{\theta}_0$ by forming $\boldsymbol{y}_i = \mathbf{A}_0^{(l)}\boldsymbol{x}_i$ and finding $\boldsymbol{\theta}_y$ as the $\boldsymbol{\theta}$ value that minimizes $\sum_{i=-m_0+1}^{0} ||\boldsymbol{y}_i - \boldsymbol{\theta}||$, using the same algorithm in Step 1. Set $\boldsymbol{\theta}_0^{(l+1)} = [\mathbf{A}_0^{(l)}]^{-1}\boldsymbol{\theta}_y$.

4. Repeat Steps 2-3 until the following condition is satisfied:

$$||\widehat{\boldsymbol{\theta}}_0^{(l)} - \widehat{\boldsymbol{\theta}}_0^{(l-1)}|| \Big/ ||\widehat{\boldsymbol{\theta}}_0^{(l-1)}|| \leq \epsilon,$$

where ϵ is a pre-specified small positive number (e.g., $\epsilon = 10^{-4}$). Then, the algorithm stops at the lth iteration and return the value of $(\widehat{\boldsymbol{\theta}}_0^{(l)}, \widehat{\mathbf{A}}_0^{(l)})$ as the final estimate for $(\boldsymbol{\theta}_0, \mathbf{A}_0)$.

The AEC dataset

[Insert Tables 1-2 around here]

References:

Bedall, K. F., and Zimmermann, H. (1979), "Algorithm AS 143. The Mediancentre," *Applied Statistics*, 28, 325–328.

Hettmansperger, T. P., and Randles, R. H. (2002), "A Practical Affine Equivariant Multivariate Median," *Biometrika*, 89, 851–860.

Randles, R. H. (2000), "A Simpler, Affine Invariant, Multivariate, Distribution-Free Sign Test," *Journal of the American Statistical Association*, 95, 1263–1268.

Runger, G. C., and Prabhu, S. S. (1996), "A Markov Chain Model for the Multivariate Exponentially Weighted Moving Averages Control Chart," *Journal of the American Statistical Association*, 91, 1701–1706.

Tyler, D. E. (1987), "A Distribution-Free M-Estimator of Multivariate Scatter," *The Annals of Statistics*, 15, 234–251.

Table 1: The AEC dataset

i	Capacitance	Dissipation	Leakage	i	Capacitance	Dissipation	Leakage
1	443	5.81	21.5	51	443	4.29	14.9
2	448	4.53	25.3	52	441	5.39	26.8
3	443	4.23	33.7	53	443	4.43	15.3
4	446	4.65	17.6	54	440	5.62	25.3
5	439	3.65	20.7	55	444	4.19	18.4
6	435	3.98	18.7	56	474	5.34	28.4
7	447	4.17	19.5	57	469	4.25	21.2
8	454	4.45	21.8	58	459	4.25	29.5
9	445	5.39	20.8	59	443	5.39	24.5
10	443	4.39	18.7	60	449	4.10	24.5
11	442	4.67	31.1	61	443	4.29	20.4
12	445	4.55	31.2	62	447	4.29	15.5
13	446	4.41	29.6	63	466	4.85	26.6
14	448	4.65	37.7	64	449	4.63	33.1
15	446	4.32	18.2	65	456	4.36	14.7
16	446	6.01	19.5	66	445	3.95	25.9
17	459	4.54	16.5	67	442	5.68	23.5
18	441	5.39	25.7	68	453	4.36	17.7
19	439	5.39	20.8	69	441	4.34	18.3
20	439	4.23	17.3	70	465	3.93	19.3
21	454	4.47	16.5	71	473	4.32	29.6
22	440	4.25	23.0	72	447	4.36	15.4
23	440	4.69	31.3	73	449	4.36	19.5
24	445	4.25	22.4	74	456	3.67	25.6
25	469	3.85	23.6	75	449	4.77	24.9
26	447	4.87	27.5	76	449	4.27	28.6
27	463	3.49	20.9	77	446	4.37	30.0
28	457	4.55	19.5	78	477	4.25	26.5
29	438	6.32	19.8	79	445	5.39	18.7
30	449	6.76	22.9	80	454	3.85	17.2
31	440	4.82	19.2	81	445	4.25	20.7
32	446	4.74	24.7	82	439	4.17	17.1
33	445	4.25	17.6	83	445	5.61	23.9
34	439	4.94	26.4	84	449	4.56	31.1
35	463	4.64	27.6	85	450	4.65	30.7
36	471	4.91	30.1	86	449	3.59	23.9
37	448	4.65	18.5	87	457	3.49	25.9
38	445	5.23	20.8	88	442	4.52	31.0
39	469	4.35	17.5	89	451	3.85	19.5
40	453	4.25	18.2	90	444	5.80	21.9
41	434	3.93	16.5	91	441	5.17	21.7
42	459	4.37	18.5	92	471	4.65	31.3
43	447	3.95	24.7	93	448	4.56	46.5
44	445	4.27	17.5	94	440	5.98	23.9
45	446	4.99	28.6	95	440	4.67	25.9
46	441	5.31	19.8	96	441	4.87	32.8
47	435	4.23	14.4	97	443	5.98	21.9
48	450	4.04	20.8	98	448	4.13	19.9
49	442	4.49	17.9	99	470	4.71	27.5
50	451	4.37	17.2	100	447	4.15	16.7

Table 2: The AEC dataset (continued)

i	Capacitance	Dissipation	Leakage	i	Capacitance	Dissipation	Leakage
101	454	4.47	18.6	151	449	4.63	31.5
102	444	4.11	24.1	152	450	4.15	19.3
103	465	3.95	45.9	153	442	4.72	28.2
104	444	3.42	25.5	154	463	4.32	18.6
105	444	3.98	17.9	155	446	5.62	19.8
106	449	4.14	19.6	156	446	4.98	23.9
107	450	4.09	21.0	157	445	3.95	19.5
108	454	3.85	16.5	158	446	5.67	20.5
109	443	4.47	36.7	159	448	4.41	31.2
110	441	5.19	21.0	160	449	4.25	16.5
111	449	4.67	27.9	161	434	3.49	17.9
112	447	3.56	15.9	162	469	4.75	31.4
113	437	4.47	21.5	163	439	4.49	15.3
114	459	4.52	26.4	164	448	4.33	27.9
115	445	4.25	16.5	165	456	3.47	27.2
116	456	4.13	20.9	166	493	4.12	19.5
117	485	3.96	27.5	167	447	3.95	21.2
118	463	4.50	16.7	168	446	4.39	30.6
119	465	4.97	26.8	169	449	3.98	20.4
120	449	4.26	20.9	170	453	4.45	18.6
121	446	4.30	30.9	171	456	4.36	15.3
122	439	5.55	23.2	172	449	4.35	18.9
123	457	4.53	25.9	173	459	4.19	27.5
124	446	4.39	30.6	174	449	4.05	25.6
125	447	4.25	15.2	175	446	5.63	21.9
126	443	4.63	17.2	176	445	4.35	18.3
127	434	4.27	15.9	177	443	4.69	26.5
128	447	4.22	24.7	178	437	4.45	15.7
129	470	4.83	30.5	179	441	4.56	35.0
130	445	3.95	16.5	180	447	3.47	19.5
131	446	4.59	36.7	181	440	4.08	20.5
132	439	5.93	25.7	182	439	5.73	24.3
133	445	4.95	33.5	183	436	3.92	17.4
134	466	4.58	28.6	184	440	4.52	15.4
135	459	4.09	17.7	185	446	5.62	19.9
136	440	4.98	21.0	186	439	4.47	16.7
137	452	4.44	37.6	187	445	4.32	19.5
138	446	4.24	15.6	188	439	4.27	20.1
139	441	5.33	19.6	189	442	4.11	21.3
140	444	4.50	30.0	190	442	4.98	29.0
141	467	3.92	18.5	191	438	3.83	14.7
142	468	4.58	34.8	192	453	4.27	17.3
143	445	4.74	20.0	193	448	4.93	17.5
144	446	4.55	21.9	194	447	4.39	19.8
145	472	4.81	27.0	195	447	4.15	30.3
146	436	5.63	29.1	196	447	4.52	17.5
147	457	4.17	17.4	197	465	4.37	24.7
148	448	4.32	29.6	198	447	4.47	17.2
149	485	3.45	25.7	199	443	4.73	31.3
150	450	4.67	29.5	200	456	4.37	16.7