

# حول مرشح كالمن الجذر التربيعي لبعض النماذج الخطية الحركية والمقارنة مع مرشح كالمن الاعتيادي

رسالة تقدّم بها

زيد طارق صالح عبّاوي الخالدي

بكالوريوس إحصاء

إلى

مجلس كلية علوم الحاسبات والرياضيات في جامعة الموصل  
وهي جزء من متطلبات نيل شهادة ماجستير علوم في الإحصاء

بإشراف

الدكتور طالب شريف جليل

أستاذ مساعد

كلية علوم الحاسبات والرياضيات – قسم الإحصاء

# بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

اقْرَأْ بِاسْمِ رَبِّكَ الَّذِي خَلَقَ ﴿١﴾ خَلَقَ الْإِنْسَانَ مِنْ عَلَقٍ ﴿٢﴾

اقْرَأْ وَرَبُّكَ الْأَكْرَمُ ﴿٣﴾ الَّذِي عَلَّمَ بِالْقَلَمِ ﴿٤﴾

عَلَّمَ الْإِنْسَانَ مَا لَمْ يَعْلَمْ ﴿٥﴾

صدق الله العظيم

## شكر وتقدير

الحمد لله رب العالمين.. والصلاة والسلام على محمد الأمين.. وعلى آله وصحبه أجمعين..

أقدم عظيم امتناني وخالص شكري للأستاذ الفاضل الدكتور طالب شريف جليل لما قدمه لي من توجيهات وتوصيات طيلة فترة إشرافه على الرسالة، كما لا أنسى أن أقر بفضلته عليّ بعد الله في زرع الرغبة لدي لخوض غمار البحث العلمي.. فجزاه الله عني خير الجزاء..

كما أقدم شكري العميق للأساتذة الأفاضل أعضاء لجنة المناقشة لتفضلهم بمناقشة رسالتي مثنياً في الوقت نفسه التصويبات والآراء المقدمة لإغناء الرسالة بأسس البحث العلمي وقواعده.. ويسرني أن أتقدم بالشكر الجزيل لأساتذتي الأكارم في قسم الإحصاء الذين كان لهم الفضل الكبير للوصول بي إلى هذه المرحلة، أدامهم الله ذخراً لهذا القسم..

وأقدم امتناني لمنتسبي المكتبة المركزية - قسم الدوريات، ومكتبة الكلية لتعاونهم الدائم، وفقهم الله لخدمة بلدنا العزيز..

ولا أنسى أن أقدم شكري واحترامي إلى الزملاء في الدراسة لما قدموه لي من دعم لوجستي لإتمام هذه الرسالة، فبارك الله فيهم جميعاً..  
وجزا الله خيراً كل من دعا لي في ظهر الغيب..

## الملخص

وضع كالمن في أوائل ستينات القرن الماضي، مجموعة من المعادلات التعاقبية على هيئة خوارزمية زمنية تعاقبية بغية إيجاد المقدر الخطي الأمثل لحالة (State) النظام الخطي الحركي. عرفت هذه الخوارزمية بمرشح كالمن الاعتيادي. إلا أن هذه الخوارزمية ليست ملائمة لجميع الظروف التي يتطلبها النظام. إن أحد أهم المشاكل التي تواجهنا عند استخدام خوارزمية مرشح كالمن الاعتيادي في التطبيقات العملية - خصوصاً التطبيقات التي تتطلب دقة عالية في النتائج- هي مشكلة فقدان مصفوفة التباين والتباين المشترك لبعض خواصها الضرورية، مثل عدم كونها أكيدة الإيجابية، أو تحولها إلى مصفوفة شاذة، ويحدث ذلك بسبب الخاصية التعاقبية لخوارزمية الترشيح الاعتيادية.

تمّ في هذه الدراسة توضيح بعض الخوارزميات البديلة لخوارزمية مرشح كالمن الاعتيادي، وتعرف هذه الخوارزميات بمرشح كالمن الجذر التربيعي، حيث تكون هذه الخوارزميات أقل تأثراً بالمشاكل العددية. وقد تمّ توضيح ثلاث خوارزميات بشكل نظري وتمّ ذلك على بعض النماذج الخطية الحركية في الأزمنة المتقطعة بالاعتماد على أسلوب بيز، وأيضاً تمّ إجراء تطبيق عملي باستخدام بيانات المحاكاة لمقارنة اثنتين من الخوارزميات الثلاث مع خوارزمية مرشح كالمن الاعتيادي، وتمّ التركيز في المقارنة على مصفوفة التباين والتباين المشترك وربحية المرشح.

## المحتويات

الصفحة	الموضوع
١	المقدمة
٥	الفصل الأول مقدمة في استدلال بيز والنماذج الخطية الحركية
٦	١-١-١- مقدمة في الاستدلال
٦	١-٢-١- مبرهنة بيز
٧	١-٣-١- طبيعة استدلال بيز
٩	١-٣-١-١- التوزيع الأولي ذو المعلومات القليلة
٩	١-٣-١-١-١- التوزيع الأولي ذو المعلومات القليلة غير القياسي
٩	١-٣-١-٢- التوزيع الأولي ذو المعلومات القليلة القياسي
١٠	١-٣-٢- التوزيع الأولي ذي المعلومات
١٠	١-٤-١- العائلات المتألفة
١٢	١-٥-١- تقدير بيز
١٣	١-٥-١-١- تقدير بيز لتعظيم دالة التوزيع اللاحق
١٤	١-٥-٢- تقدير الإمكان الأعظم
١٥	١-٥-٣- تقدير بيز ذو أقل متوسط مربع خطأ
١٦	١-٥-٤- تقدير بيز ذو أقل قيمة متوسط خطأ مطلق
١٧	١-٦-١- تحليل بيز المتسلسل
١٧	١-٧-١- النماذج الخطية الحركية
١٩	١-٨-١- مبرهنة بيز والنماذج الخطية الحركية
٢٣	١-٩-١- بعض النماذج الخطية الحركية
٢٤	١-٩-١-١- النموذج الدرجي
٢٥	١-٩-٢- نموذج النمو الخطي الحركي المركب
٢٧	الفصل الثاني مرشح كالمن ومشكلة التباعد في مصفوفة التباين والتباين المشترك
٢٨	١-٢-١- مقدمة
٣١	١-٢-٢- بعض التطبيقات التي يستخدم فيها مرشح كالمن
٣٢	١-٣-٢- مرشح كالمن ومبرهنة بيز
٣٦	١-٤-٢- أنواع مرشح كالمن

٣٧	٥-٢- الصيغة القطرية للمصفوفة
٣٧	٦-٢- المبرهنة الطيفية
٣٨	٧-٢- الجذر التربيعي للمصفوفة
٣٨	١-٧-٢- الجذر التربيعي للمصفوفة باستخدام التحليل الطيفي
٣٩	٢-٧-٢- تحليل چولسكي
٤١	٨-٢- مرشح كالمن الجذر التربيعي
٤١	١-٨-٢- التباعد
٤٣	٢-٨-٢- مرشح الجذر التربيعي لبوتر
٤٦	٣-٨-٢- مرشح كالمن-شميدت الجذر التربيعي
٤٩	٤-٨-٢- مرشح الجذر التربيعي لمصفوفة التباين والتباين المشترك
٥٢	<b>الفصل الثالث الجانب التطبيقي</b>
٥٣	١-٣- مقدمة
٥٣	٢-٣- التطبيق لمرشح الجذر التربيعي لبوتر
٦٠	٣-٣- التطبيق لمرشح الجذر التربيعي لمصفوفة التباين والتباين المشترك
٦٦	<b>الاستنتاجات والتوصيات</b>
٦٧	<b>المصادر</b>
٧٠	<b>الملاحق</b>

## قائمة بالأشكال الواردة في الرسالة

الصفحة	عنوان الشكل	رقم الشكل
٩	منحنى دالة الإمكان المهيمنة على التوزيع الأولي	(١-١)
١٤	دالة الكلفة المنتظمة ومنحنى التوزيع اللاحق	(٢-١)
١٦	منحنى التوزيع المتماثل إضافة إلى الوسط والوسيط والمنوال	(٣-١)
٢٣	منحنى التوزيعين الأولي واللاحق عند الزمن ( $t=1$ )	(٤-١)
٢٣	منحنى التوزيعين الأولي واللاحق عند الزمن ( $t=2$ )	(٥-١)
٣٠	الفرق بين التمهيد والترشيح والتنبيؤ	(١-٢)
٣١	خطوات الخوارزمية التعاقبية التي وضعها كالمن لإيجاد المقدر الأمثل	(٢-٢)
٤٠	المراحل المتعاقبة لإيجاد المصفوفة المثلثية السفلى لجولسكي	(٣-٢)

## المقدمة

إن الغاية الأساسية من الترشيح هي إيجاد المقدّر الأمثل (Optimal Estimator) للمعلمة. حيث أن مصطلح الترشيح يعني في علم الإحصاء التقدير في الزمن الحاضر، أي إيجاد مقدّر أمثل للمعلمة ( $\theta_r$ ) عندما تكون ( $\tau = t$ ) حيث أن ( $t$ ) تمثل الزمن الحاضر [Kalman, 1960]. وكانت الأساسيات التي بنيت عليها نظرية التقدير عند ظهورها للمرة الأولى تعتمد على جعل دالة تباين الخطأ أقل ما يمكن، وكما هو معلوم فإن طريقة المربعات الصغرى وجدت لأول مرة على يد العالم (Gauss) عام 1795. ومنذ ذلك الحين تتابعت الدراسات الموسعة في هذا المجال، وخصوصاً فيما يتعلق بالتقدير الخطي الأمثل (Optimal Linear Estimation) وقد ظهرت أولى الدراسات الرائدة في مسألة الترشيح في مطلع الأربعينات من القرن الماضي، حيث قام كلٌّ من (Kolmogorov) عام 1941، و(Wiener) عام 1942 بشكل مستقل بتطوير تقنية لإيجاد مقدّر خطي ذي أقل متوسط مربع خطأ (Linear Minimum Mean-Square Error Estimator) والذي لقي اهتماماً بالغاً وكان له الأثر الكبير لاحقاً في تطوير فكرة مرشح كالمن. [Sorenson, 1970]، ثم توجّج [Wiener, 1949] دراسته تلك بنتائج أخرى تمخضت عن وضع خوارزمية تعاقبية (Recursive Algorithm) لإيجاد المقدّر الخطي الأمثل، وكانت هذه الخوارزمية مقيدة بشروط منها أن المشاهدة المدروسة ذات بُعد أحادي (لا توجيهي Scalar) فضلاً عن المعلمة، وأن البيانات شبه غير منتهية ( $t_0 \rightarrow -\infty$ ) فضلاً عن ذلك فإن العملية تتسم بالاستقرار (Stationary Process).

وفي عام 1950 قام (Zadeh & Ragazzini) بتطوير النتائج التي توصل إليها (Wiener) وذلك بجعلها أكثر عمومية وذلك من حيث جعل البيانات منتهية ( $t_0 > -\infty$ ) ولكن العمليات تتسم بعدم الاستقرار (Non-Stationary Processes). تلا ذلك دراسات أخرى في المجال نفسه، نذكر منها على سبيل المثال لا الحصر الدراسة التي قام بها (Dolph & Woodbury) في عام 1952 و (Darlington) في عام 1959 والتي لم تتعد النتائج التي توصل إليها (Zadeh & Ragazzini). [Kailath, 1974]

ومع كثرة الدراسات في هذا المجال فقد بقيت عقيمة نوعاً ما عن تفسيرها لظواهر عديدة ظهرت خصوصاً في أواخر الخمسينات بعد دخول عصر الفضاء وظهور مشكلة تحديد مدار القمر الصناعي (Satellite Orbit) من حيث السرعة (Velocity) كمعلمة أولى والتعجيل (Acceleration) كمعلمة ثانية، وذلك لأسباب عدة، مما حدا بالباحثين إلى محاولة إيجاد خوارزمية تتغلب على هذه المشاكل.



وكانت أولى الخوارزميات التي تعالج هذه المشاكل هي الخوارزمية التي وضعها (Swerling) عام (1959)، حيث تمكن سويرلنك من وضع أول خوارزمية تعاقبية لحل المشاكل التي تعالج مشكلة تعدد المعلمات، فضلاً عن المشاكل الأخرى، بعد ذلك قام [Kalman, 1960] بوضع خوارزمية تعاقبية أكثر تعقيداً من تلك التي وضعها سويرلنك لكنها تميزت عنها بكونها أول خوارزمية تهتم بدراسة مشاكل تقدير الحالة الحركي Dynamical State Estimation Problems، ويسرت هذه الخوارزمية فهم العديد من المشاكل والظواهر التطبيقية وتحليلها، وقد كانت الخوارزمية التي وضعها كالمن تهتم بدراسة تقدير الحالة الحركي في الأزمنة المتقطعة (Discrete Time) أي أن  $t=1,2,\dots$ . ثم قام [Kalman & Bucy, 1961] بوضع خوارزمية لتقدير الحالة الحركي في الأزمنة المستمرة (Continuous Time) وهو ما عرف لاحقاً بـ (Kalman-Bucy Filter). [Kailath, 1974]

وقد اعتمد كالمن في إيجاده للمقدر الخطي الأمثل على فكرة الإسقاط المتعامد (Orthogonal Projection)، وقد توصل كالمن إلى أن المقدر الخطي الأمثل للمعلمة  $(\theta_t)$  هو التوقع الشرطي لـ  $(\theta_t)$  مشروطاً بكافة المعلومات المتوافرة حتى الزمن  $(t)$  أي:

$$E(\theta_t | D_t)$$

حيث أن  $D_t$  تمثل المعلومات المتوافرة حول المعلمة حتى الزمن الحاضر  $(t)$ .  
 وأن هذا المقدر يمكن الحصول عليه إذا توافر أحد الشرطين الآتيين:  
 ١. إذا توزع كلٌّ من المعلمة  $(\theta_t)$  والمشاهدة  $(y_t)$  وفق التوزيع الطبيعي.  
 ٢. إذا كانت دالة الخسارة هي دالة الخسارة التربيعية (Quadratic Loss Function)، أي

$$\text{أن: } l(\hat{\theta}_t, \theta_t) = (\hat{\theta}_t - \theta_t)^2 \text{ [Kalman, 1960]}$$

ومن ثم لاحظ كل من هارسون وستيفنس [Harrison & Stevens, 1976]، أن الشروط أعلاه مماثلة لإيجاد مقدر بيز ذي أقل متوسط مربع خطأ Minimum Mean Square Error Estimator أو ما يعرف بالمقدر الأمثل والذي يساوي التوقع المشروط لـ  $(\theta_t)$  بالمعلومات  $(D_t)$ ، وأثبتنا أيضاً أن مقدر بيز ذا أقل متوسط مربع خطأ عند إيجاده بشكل متعاقب (Sequential) فإنه يتساوى مع المقدر الخطي الأمثل الذي وضعه كالمن. إن استخدام نظرية بيز في عملية تكوين مرشح كالمن تعطي سلسلة وسهولة عاليتين في العمليات الحسابية التي يتطلبها تكوين مرشح كالمن مقارنة مع الأسلوب الذي وضعه كالمن للمرة الأولى عام 1960 والذي يتسم بالتعقيد بعض الشيء. ومن ابرز من استخدم أسلوب بيز في تكوين مرشح كالمن [Maybeck, 1979] و [Meinhold & Singpurwalla, 1983] و [West & Harrison, 1989] و [Calder, et al. 2003] وغيرهم آخرون.

بعد ذلك تتابعت الدراسات النظرية والتطبيقية وبشكل موسع جداً على ما سمي لاحقاً بنظرية مرشح كالمن (Kalman Filter Theory)، تضمنت تطوير خوارزميات محدثة لخوارزمية مرشح كالمن الاعتيادي (Ordinary Kalman Filter Algorithm) وذلك للتغلب على المشاكل التي تواجه المستخدم لهذه الخوارزمية، مثل مشكلة النماذج غير الخطية، أو عندما يكون تباين الأخطاء غير معلوم، أو عندما تكون الأخطاء العشوائية مترابطة، انظر [الحمداي، 1996]، وكذلك الخوارزميات التي وضعت للتغلب على المشاكل العددية التي تحصل في مصفوفة التباين  $C_t^*$  بسبب الإجراء التعاقبي.

إن المشاكل التي تحدث في مصفوفة التباين تؤدي إلى حصول تباعد (Divergence) في خوارزمية مرشح كالمن الاعتيادي، مما يؤدي إلى فشل الإجراء، أو الحصول على نتائج غير دقيقة أو غير واقعية كأن ينخفض التباين بشكل ملحوظ، أو بقاء الخوارزمية غير مستقرة أي عدم الوصول إلى حالة الاستقرار (Steady State). لذلك وضعت صيغ وخوارزميات عدة للتغلب على مشكلة التباعد هذه.

وكانت معظم تلك الخوارزميات تعتمد على فكرة إيجاد مصفوفة الجذر التربيعي لمصفوفة التباين وذلك لأن الدقة المتوافرة تكون أعلى، فضلاً عن أن مصفوفة الجذر التربيعي تكون أكيدة الإيجابية. [Bierman, 1977]، لذلك عرفت هذه الخوارزميات بشكل عام بخوارزمية مرشح كالمن الجذر التربيعي (Square Root Kalman Filter). [Maybeck, 1979].

ويعد بوتز (Potter) أول من وضع خوارزمية مرشح كالمن الجذر التربيعي في عام (1963)، عرفت خوارزميته تلك بخوارزمية مرشح الجذر التربيعي لبوتز [Carlson, 1973]، وتشتت هذه الخوارزمية أن تكون المشاهدات مفردة البعد (Scalar) أي أن  $(m=1)$ ، وأن تكون المعلمات خالية من التشويش (Noise) أي تكون  $(W_t = 0)$ ، مما جعلها مقيدة بالرغم من كفاءتها.

ثم قدم [Bellantoni, & Dodge, 1967] بحثاً تضمن نتائج جديدة في هذا المجال، حيث وضع خوارزمية جديدة لا تنقيد بشروطي خوارزمية بوتز. مع ذلك فإن هذه الخوارزمية تكون غير كفوءة بالرغم من فعاليتها وذلك في الحالات التي تكون فيها عدد المشاهدات أقل من عدد المعلمات في النموذج المدروس، أي أن  $(m < n)$ . [Kaminski, et al. 1971].

وقدم [Andrews, 1968] خوارزمية أخرى لمعالجة مشاكل مصفوفة التباين، حيث تعتمد هذه الخوارزمية في إيجاد مصفوفة الجذر التربيعي على تحليل چولسكي Cholesky Decomposition. تلا ذلك العديد من الدراسات والبحوث التي تمخضت عن تطوير خوارزميات جديدة لمرشح كالمن الجذر التربيعي تعالج بعض المشاكل العامة، مثل الخوارزمية التي وضعها [Carlson, 1973] والتي تستخدم في النماذج التي تحوي على أخطاء مترابطة (Correlated Errors).

\* المقصود مصفوفة التباين والتباين المشترك (Covariance Matrix)، وقد كتبت بهذه الصيغة في الرسالة لغرض الاختصار .

أو الخوارزميات التي وضعت لمعالجة مشكلة النماذج غير الخطية مثل خوارزمية مرشح كالمن الجذر التربيعي الموسع (Extended Square Root Kalman Filter)، أو خوارزمية الجذر التربيعي المتوازية (Parallel Square Root Algorithm). [Lu, et al. 1992].

لقد استخدم مرشح كالمن الجذر التربيعي على نطاق واسع في الدراسات والتطبيقات الهندسية وخصوصاً هندسة الاتصالات مثل تطبيقات أجهزة الرصد الجوي والفضائي، وفي الصناعات الدوائية، حيث تتطلب هذه التطبيقات دقة عالية في الحسابات.

ولقد وجهنا اهتمامنا في هذه الرسالة بدراسة مرشح كالمن الجذر التربيعي الذي يعد أحد الطرائق لمعالجة المشاكل التي تظهر في مصفوفة التباين في مرشح كالمن الاعتيادي نتيجة للتكرار المتعاقب، حيث تم توضيح بعض خوارزميات مرشح الجذر التربيعي فضلاً عن توضيح الخوارزمية الأصلية لكالمن وذلك في بعض النماذج الخطية الحركية، وذلك بالاعتماد على أسلوب بيز، وتم إجراء تطبيق عملي على بيانات المحاكاة Simulation Data التي تم توليدها باستخدام نظام ماتلاب Matlab 7.0، وقد أُدّت هذه البيانات لتكون متوافقة مع النماذج المدروسة مع مراعاة متطلبات الدراسة، وتمت مقارنة النتائج بين مرشح كالمن الجذر التربيعي مع مرشح كالمن الاعتيادي وذلك لبيان الفرق بين الإجراءين.

ولإتمام الهدف من الرسالة بشكل علمي، ولكي يتم الإلمام بمفهوم المواضيع الواردة فيها، فقد جاءت الرسالة بثلاثة فصول، تناول الفصل الأول استدلال بيز ونظرية التقدير والنماذج الخطية الحركية، حيث تم توضيح طبيعة استدلال بيز، فضلاً عن بعض المقدرات البيزية المعتمدة على دالة المخاطرة، كما تم توضيح العلاقة بين استدلال بيز والنماذج الخطية الحركية، وتضمن الفصل أيضاً شرحاً مبسطاً لبعض النماذج الخطية الحركية فضلاً عن إيجاد التوزيعين السابق واللاحق لمعاملات تلك النماذج. فيما خصص الفصل الثاني للبحث في نظرية الترشيح وتاريخ تطورها وتوضيح أهم خوارزمية في هذا المجال، وهي خوارزمية مرشح كالمن الاعتيادي، وتم أيضاً توضيح مشكلة التباين التي تحصل في مصفوفة التباين نتيجة للإجراء التكراري المتعاقب وكيفية التغلب على هذه المشكلة، حيث تم توضيح ثلاث خوارزميات لمرشح كالمن الجذر التربيعي وهن خوارزمية مرشح الجذر التربيعي لبوتر، وخوارزمية مرشح كالمن-شميدت الجذر التربيعي، وخوارزمية مرشح الجذر التربيعي لمصفوفة التباين والتباين المشترك، فضلاً عن بعض الفقرات الأخرى المهمة. واختتمت الرسالة بفصل ثالث تم فيه إجراء تطبيق عملي باستخدام بيانات المحاكاة لاثنتين من خوارزميات مرشح الجذر التربيعي، فضلاً عن مقارنة النتائج مع نتائج خوارزمية مرشح كالمن الاعتيادي.

# الفصل الأول

مقدمة في استدلال بيز  
والنماذج الخطية الحركية

*Introduction in Bayesian Inference  
& The Dynamic Linear Models*

## الفصل الأول

### مقدمة في استدلال بيز والنماذج الخطية الحركية

#### ١-١- مقدمة في الاستدلال:

للتعرف على استدلال بيز لا بد من أخذ فكرة عن ماهية الاستدلال. حيث يعرف الاستدلال بأنه أحد فروع علم الإحصاء والذي يهتم بجمع المعلومات حول معلمات التوزيعات، ومن أساليبه في ذلك نظرية التقدير واختبار الفرضيات.

وقد شهد القرن العشرون تطوراً متسارعاً في الأساليب (approaches) المختلفة المستخدمة في الاستدلال الإحصائي، حيث يمكن تصنيف الاستدلال الإحصائي إلى مدرستين رئيسيتين تبعاً لاختلاف وجهة النظر الإحصائية حول ماهية المعلمة، وهاتان المدرستان هما:

١. المدرسة الكلاسيكية أو ما يسمى بمدرسة المعاينة (Classical or Sampling School)، وتفترض هذه المدرسة عند دراستها للمعلمة بأنها -أي المعلمة- ثابت غير معلوم (Unknown Constant)، وتستخلص هذه المدرسة دراستها حول المعلمة من خلال المعلومات التي توفرها العينة، لذا تدعى بمدرسة المعاينة. وممن برز في هذه المدرسة العالم فيشر ونيمان وبيرسون وغيرهم آخرون.

٢. مدرسة بيز (Bayesian School) نسبة إلى الإحصائي البريطاني توماس بيز، حيث تفترض هذه المدرسة بأن المعلمة متغير عشوائي له توزيع احتمالي. وتعتمد هذه المدرسة في دراستها للمعلمة على المعلومات التي توفرها العينة فضلاً عن المعلومات التي تأتي من الخبرة أو الاعتقاد الشخصي (Personal Believes).

#### ٢-١- مبرهنة بيز (Bayes theorem):

توماس بيز (Thomas Bayes 1702-1761) هو راهب بريطاني ساهم في إغناء نظرية الاحتمال ببحثين نشرهما بعد وفاته زميله ريتشارد برايس (Richard Price) في عامي (١٧٦٣ و ١٧٦٤)، وقد أضاف إلى البحثين بعض الأمثلة والتعليقات فضلاً عن المقدمة.

[Barnett, 1982]

ومن أبرز الذين ساهموا في تطوير أسلوب بيز في المجالين النظري والتطبيقي العلماء (Laplace) و (Lindly) و (Box & Tiao) و (Zellner) و (Jeffrey) وغيرهم آخرون.

وبالإمكان الحصول على صيغة مبرهنة بيز الاحتمالية من المفهوم العام لاحتمال الشرطي، وكالاتي:

على فرض أن المتغير العشوائي المنقطع  $X$  يأخذ القيم  $(x_i's ; i = 1, 2, \dots, n)$ ، ولنفرض أيضاً أن  $Y$  متغير عشوائي منقطع أيضاً ويأخذ القيم  $(y_j's ; j = 1, 2, \dots, m)$ ، فعندئذ تكون دالة الكتلة الاحتمالية المشتركة (Joint Probability Mass Function) لهذين المتغيرين هي  $P(X = x_i, Y = y_j)$ ، حيث يمكن كتابة هذه الدالة بأسلوب آخر وذلك بالاعتماد على قانون ضرب الاحتمالات:

$$P(x_i, y_j) = \begin{cases} P(x_i | y_j) P(y_j) \\ P(y_j | x_i) P(x_i) \end{cases} \dots\dots\dots(1.1)$$

حيث أن  $P(x_i | y_j)$  يمثل الاحتمال الشرطي لـ  $x$  عندما يعطى  $y$ ، ومنه نحصل على:

$$P(x_i | y_j) = \frac{P(x_i, y_j)}{P(y_j)}$$

أما  $P(y_j)$  فهي تمثل دالة الكتلة الاحتمالية الحدية للمتغير  $Y$ ، ويتم إيجادها من قانون الاحتمال الكلي:

$$P(y_j) = \sum_{i=1}^n P(y_j | x_i) P(x_i) \dots\dots\dots(1.2)$$

وبالتعويض بالمعادلتين (1,1) و(1,2) نحصل على الصيغة العامة لمبرهنة بيز [Vaseghi, (2000)]:

$$P(x_i | y_j) = \frac{P(y_j | x_i) P(x_i)}{\sum_{i=1}^n P(y_j | x_i) P(x_i)}$$

وبالطريقة نفسها يمكن الحصول على الصيغة العامة لمبرهنة بيز في حالة المتغيرات المستمرة:

$$P(x | y) = \frac{P(y | x) P(x)}{\int_{\Omega_X} P(y | x) P(x) dx}$$

حيث أن  $P(\cdot)$  تمثل دالة الكثافة الاحتمالية (Probability Density Function).

### ٣-١- طبيعة استدلال بيز (Nature of Bayesian Inference)

افرض أن  $(y_1, y_2, \dots, y_n)$  متغيرات عشوائية لها توزيع احتمالي  $P(y | \theta)$ ، حيث  $\theta$  تمثل معلمة التوزيع، فعندئذ:

$$P(\theta | y_1, y_2, \dots, y_n) = \frac{P(y_1, y_2, \dots, y_n | \theta) P(\theta)}{P(y_1, y_2, \dots, y_n)}$$

حيث أن:

$$P(y_1, y_2, \dots, y_n) = c^{-1} = \left\{ \begin{array}{l} \sum_{p=1}^k P(y_1, y_2, \dots, y_n | \theta_p) P(\theta_p) \\ \int_{\theta \in \Omega} P(y_1, y_2, \dots, y_n | \theta) P(\theta) d\theta \end{array} \right\} \begin{array}{l} \theta \text{ متغير متقطع} \\ \theta \text{ متغير مستمر} \end{array}$$

إذن:

$$P(\theta | y_1, y_2, \dots, y_n) = c P(y_1, y_2, \dots, y_n | \theta) P(\theta)$$

وبما أن  $c$  قيمة ثابتة (ثابت التناسب) - لأنها خالية من  $\theta$  - فيمكن كتابة صيغة مبرهنة بيز بالصيغة الآتية:

$$P(\theta | y_1, y_2, \dots, y_n) \propto P(y_1, y_2, \dots, y_n | \theta) P(\theta) \quad \dots\dots\dots(1.3)$$

وكما نرى فإن الصيغة العامة تتكون من ثلاثة أجزاء، حيث يتكون الطرف الأيمن من عاملين الأول  $P(\theta)$  يمثل المعلومات المتوافرة حول المعلمة (خبرة أو اعتقاد شخصي) قبل أخذ أي مشاهدة، ويدعى بالاحتمال الابتدائي أو التوزيع الأولي (Prior Distribution) للمعلمة وهو يلعب دوراً مهماً في مبرهنة بيز ويصنف إلى صنفين سيجري توضيح كل منهما على حدا لاحقاً.

والجزء الثاني هو  $P(y_1, y_2, \dots, y_n | \theta)$  دالة الإمكان (Likelihood Function)، وتعد دالة الإمكان دالة لـ  $(\theta)$  لأننا نهتم بدراسة  $(\theta)$  ولذلك تكتب  $L(\theta)$ ، وهي تمثل المعلومات المتوافرة من البيانات حول المعلمة. [Box & Tiao (1973)]

وإذا أخذنا بنظر الاعتبار المعلومات المأخوذة من البيانات فضلاً عن تلك المتوافرة من التوزيع الأولي نحصل على ما يسمى بالتوزيع اللاحق (Posterior Distribution) وهو الجزء الثالث من مبرهنة بيز، ويكتب  $P(\theta | y_1, y_2, \dots, y_n)$  ويمثل الطرف الأيسر من المعادلة (1.3). ويجب أن نؤكد هنا أن الثابت  $c$  يمثل ثابت التطبيع (Normalization Constant) وهو العامل الذي يجعل تكامل التوزيع اللاحق (أو المجموع) مساوٍ لواحد. وبذلك يمكننا كتابة مبرهنة بيز بالصيغة الآتية:

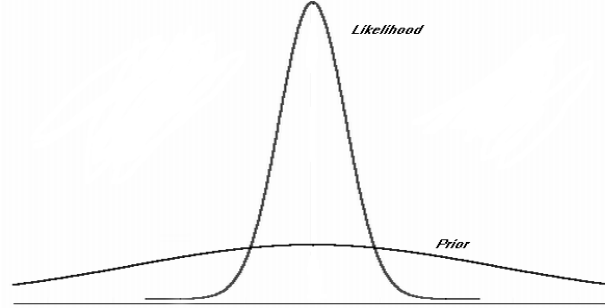
$$P(\theta | y_1, y_2, \dots, y_n) \propto L(\theta) P(\theta)$$

وبالرجوع إلى التوزيع الأولي فإنه بالإمكان تصنيفه إلى نوعين وذلك تبعاً لأهمية المعلومات التي يوفرها حول المعلمة  $\theta$ ، وهذان النوعان هما:

### ١-٣-١- التوزيع الأولي ذو المعلومات القليلة:

#### **Non-Informative Prior Distribution:**

عندما تكون دالة الإمكان مهيمنة على التوزيع الأولي فعندئذٍ يعرف ذلك التوزيع بأنه توزيع أولي ذو معلومات قليلة، ويكون هذا التوزيع ذا تأثير قليل على المنطقة التي تكون فيها دالة الإمكان ذات تأثير مدرك (Appreciable) في التوزيع اللاحق بالإضافة إلى ذلك فهو لا يأخذ قيمة كبيرة خارج تلك المنطقة - أنظر الشكل (١-١) - إن التوزيع الذي له الخواص أعلاه يعرف بتوزيع أولي منتظم موضعياً (Locally Uniform Prior Distribution) أو توزيع أولي غير كامل (Improper Prior Distribution). [Box & Tiao, 1973]



الشكل (١-١)

منحنى دالة الإمكان المهيمنة على التوزيع الأولي. [Box & Tiao, 1973]

ويقسم التوزيع الأولي ذو المعلومات القليلة إلى قسمين:

### ١-٣-١-١- التوزيع الأولي ذو المعلومات القليلة غير القياسي:

#### **Non-Standard Non-Informative Prior Distribution:**

هذا النوع من التوزيع الأولي يختلف من شخص إلى آخر، وذلك حسب الخبرة والاعتقاد الشخصي. [الحمداني، ١٩٩٦]

### ١-٣-١-٢- التوزيع الأولي ذو المعلومات القليلة القياسي:

#### **Standard Non-Informative Prior Distribution:**

في هذه الحالة يمكن إيجاد التوزيع الأولي باستخدام قانون جفريز نسبة إلى الإحصائي البريطاني (Harold Jeffreys)، حيث ينص هذا القانون على أن التوزيع الأولي لمعلمة واحدة يتناسب مع الجذر التربيعي لقياس معلومات فيشر (Square Root of Fisher Information Measure)، أي أن:



$$P(\theta) \propto (I_y(\theta))^{1/2}$$

حيث أن  $I_y(\theta)$  تمثل معلومات فيشر حول المعلمة  $\theta$  التي يوفرها المتغير العشوائي  $(y)$ . ويمكن إيجاد معلومات فيشر كالاتي:

$$I_y(\theta) = E \left[ \frac{\partial \ln f(y|\theta)}{\partial \theta} \right]^2$$

كما يمكن إيجادها بصيغة بديلة. [Kendall & Stuart, 1976]:

$$I_y(\theta) = - E \left[ \frac{\partial^2 \ln f(y|\theta)}{\partial \theta^2} \right]$$

### ١-٣-٢- التوزيع الأولي ذي المعلومات (Informative Prior Distribution):

هو ذلك التوزيع الذي لا تكون دالة الإمكان مهيمنة عليه، ويكون له تأثير ملحوظ على التوزيع اللاحق. ويبدو من الواضح هنا بأنه إذا كان التوزيع الأولي مهيمناً على دالة الإمكان فهذا يعني أن القسم الأكبر من المعلومات تأتي من هذا التوزيع ولهذا يطلق على التوزيع الأولي في هذه الحالة بذوي المعلومات (Informative).

### ١-٤- العائلات المتآلفة (Conjugate Families):

إذا كان كلٌّ من التوزيع الأولي واللاحق ينتميان إلى نفس العائلة من التوزيعات فعندئذٍ يطلق على هذه التوزيعات بالتوزيعات المتآلفة. لمزيد من المعلومات أنظر [Lindgren, 1962]

◀ مثال:

على فرض أن  $(y_1, y_2, \dots, y_n)$  عينة عشوائية من مجتمع له توزيع طبيعي بمتوسط  $\theta$  وتباين معلوم  $\sigma^2$ ، أي أن:

$$f(y_i | \theta, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{1}{2\sigma^2} (y_i - \theta)^2 \quad \begin{array}{l} i=1, 2, \dots, n \\ -\infty < y_i < \infty \\ -\infty < \theta < \infty \\ \sigma^2 > 0 \end{array}$$

$$\therefore L(\theta) = \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \theta)^2$$

$$\therefore L(\theta) \propto \exp - \frac{n}{2\sigma^2} (\theta - \bar{y})^2 \quad , \bar{y} = \frac{\sum_{i=1}^n y_i}{n}$$

وافرض أيضاً أن  $\theta$  لها توزيع أولي طبيعي بمتوسط  $\theta_0$  وتباين  $\sigma_0^2$ ، أي أن:

$$P(\theta) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp - \frac{1}{2\sigma_0^2} (\theta - \theta_0)^2$$

وعليه يمكن إيجاد التوزيع اللاحق كالاتي:

$$P(\theta | y_1, y_2, \dots, y_n) \propto P(\theta) L(\theta) \\ \propto \exp - \frac{1}{2\sigma_0^2} (\theta - \theta_0)^2 \times \exp - \frac{n}{2\sigma^2} (\theta - \bar{y})^2$$

وبفتح الأقواس وحذف الحدود التي لا تحتوي على  $\theta$  نحصل على:

$$P(\theta | y_1, y_2, \dots, y_n) \propto \exp \left[ - \frac{\sigma^2 + n\sigma_0^2}{2\sigma_0^2\sigma^2} \left( \theta - \frac{\theta_0\sigma^2 + n\bar{y}\sigma_0^2}{n\sigma_0^2 + \sigma^2} \right)^2 \right]$$

والأخير يمثل نواة (Kernel) التوزيع الطبيعي بمتوسط  $\mu$ :

$$\mu = \frac{n\bar{y}\sigma_0^2 + \theta_0\sigma^2}{n\sigma_0^2 + \sigma^2}$$

$$\mu = \frac{(n\bar{y}/\sigma^2) + (\theta_0/\sigma_0^2)}{(n/\sigma^2) + (1/\sigma_0^2)} \quad \text{وبقسمة البسط والمقام على } (\sigma^2\sigma_0^2) \text{ فإن:}$$

وتباين التوزيع اللاحق  $\nu$ :

$$\nu = \frac{\sigma_0^2\sigma^2}{\sigma^2 + n\sigma_0^2}$$

وبذلك فإن التوزيع اللاحق لمعدل التوزيع الطبيعي  $\theta$  يمكن كتابته بالصيغة العامة للتوزيع الطبيعي:

$$P(\theta | y_1, y_2, \dots, y_n) = \frac{1}{\sqrt{2\pi\nu}} \exp - \frac{1}{2\nu} (\theta - \mu)^2$$

إذن كما هو ملاحظ فإن التوزيع الطبيعي الأولي بمتوسط  $\theta_0$  وتباين معلوم  $\sigma_0^2$  أعطى توزيعاً لاحقاً طبيعياً أيضاً، أي أن التوزيعان الأولي واللاحق هما من عائلة متألّفة.

<sup>1</sup> تعرف نواة التوزيع في استدلال بيز (Kernel) بأنها ذلك الجزء من دالة التوزيع الاحتمالي الذي يحتوي على المعلمة المدروسة.

■ ملاحظة:

من الأفضل كتابة متوسط وتباين التوزيع اللاحق بدلالة الدقة (Precision)، حيث دقة العينة

هي  $\pi_s = n/\sigma^2$ ، ودقة التوزيع الأولي هي  $\pi_0 = 1/\sigma_0^2$ ، أي أن:

$$\mu = \frac{\pi_s \bar{y} + \pi_0 \theta_0}{\pi_s + \pi_0} \quad \nu = \frac{1}{\pi_s + \pi_0} = \frac{1}{\pi_p}$$

ويلاحظ أن متوسط التوزيع اللاحق هو المتوسط الموزون لمعدل العينة ومعدل التوزيع الأولي، حيث تكون الأوزان هي الدقة في كل حالة. ويلاحظ أيضاً أن الدقة في التوزيع اللاحق هي الدقة في العينة مضافاً إليها الدقة في التوزيع الأولي  $\pi_p = \pi_s + \pi_0$ . والجدول (1-1) يوضح بعض الأمثلة عن العائلات المتألفة.

الجدول (1-1)

بعض التوزيعات ذات العائلات المتألفة. [Lindgren, 1962]

Likelihood Fun.	Prior Dist.	Posterior Dist.
$Bernoulli(\theta)$ $\theta^{\sum y_i} (1-\theta)^{n-\sum y_i}$	$Beta(\alpha, \beta)$ $\theta^{\alpha-1} (1-\theta)^{\beta-1}$	$Beta(\alpha + \sum y_i, \beta + n + \sum y_i)$ $\theta^{(\alpha+\sum y_i)-1} (1-\theta)^{(\beta+n+\sum y_i)-1}$
$Poisson(\theta)$ $e^{-n\theta} \theta^{\sum y_i}$	$Gamma(\alpha, \beta)$ $\theta^{\alpha-1} e^{-\beta\theta}$	$Gamma(\alpha + \sum y_i, \beta + n)$ $\theta^{(\alpha+\sum y_i)-1} e^{-(\beta+n)\theta}$
$Exponential(\theta)$ $\theta^n e^{-\theta \sum y_i}$	$Gamma(\alpha, \beta)$ $\theta^{\alpha-1} e^{-\beta\theta}$	$Gamma(\alpha + n, \beta + \sum y_i)$ $\theta^{(\alpha+n)-1} e^{-(\beta+\sum y_i)\theta}$
$N(\theta, \sigma^2) \sigma^2 \text{ Known}$ $e^{-\frac{n}{2\sigma^2}(\theta-\bar{y})^2}$	$N(\theta_0, \sigma_0^2)$ $e^{-\frac{1}{2\sigma_0^2}(\theta-\theta_0)^2}$	$N(\mu, \nu)$ $e^{-\frac{1}{2\nu}(\theta-\mu)^2}$

### 1-5- تقدير بيز (Bayes Estimation):

إن الغاية من إيجاد التوزيع اللاحق في مبرهنة بيز هي الحصول على تقدير للمعلمة غني بالمعلومات المعبرة عن واقع حال المعلمة، سواءً كان ذلك التقدير بنقطة أم بفترة. وسوف يكون اهتمامنا في هذه الدراسة حول تقدير بيز بنقطة (Bayesian Point Estimation).

وهناك عدة طرائق لإيجاد مقدر بيز النقطي، حيث تعتمد هذه الطرائق في غالبيتها على إيجادها لمقدر يجعل دالة المخاطرة الشرطية (Conditional Risk Function) أقل ما يمكن، حيث أن دالة

المخاطرة الشرطية  $\mathfrak{R}(\hat{\theta} | y_1, y_2, \dots, y_n)$ ، هي عبارة عن التوقع الشرطي لدالة كلفة الخطأ (Cost of Error Function  $C(\hat{\theta}, \theta)$ )، أو دالة الخسارة (Loss Function  $l(\hat{\theta}, \theta)$ )، أي أن:

$$\begin{aligned}\mathfrak{R}(\hat{\theta} | y) &= E[C(\hat{\theta}, \theta) | y_1, y_2, \dots, y_n] \\ &= \int_{\theta \in \Omega} C(\hat{\theta}, \theta) P(\theta | y_1, y_2, \dots, y_n) d\theta\end{aligned}$$

أي أن:

$$\begin{aligned}\hat{\theta}_B &= \arg \min \mathfrak{R}(\hat{\theta} | y_1, y_2, \dots, y_n) \\ &= \arg \min \left[ \int_{\theta \in \Omega} C(\hat{\theta}, \theta) P(\theta | y_1, y_2, \dots, y_n) d\theta \right]\end{aligned}$$

وتعرف دالة المخاطرة الشرطية بدالة مخاطرة بيز Bayesian Risk Function. ومن هذه الطرائق:

#### ١-٥-١ - تقدير بيز لتعظيم دالة التوزيع اللاحق: Maximum A Posteriori Estimation

هو المقدر الذي يعظم دالة التوزيع اللاحق  $P(\theta | y_1, y_2, \dots, y_n)$ ، ويرمز له بـ  $\hat{\theta}_{MAP}$  ويمكن إيجاده فقط عندما تكون دالة الكلفة منتظمة (Uniform Cost Function) - قيمة ثابتة لفترات معينة من قيم  $\theta$  - [Vaseghi, 2000]، وتعرف:

$$C(\hat{\theta}, \theta) = \begin{cases} 0 & \text{if } |\theta - \hat{\theta}| < \varepsilon \\ 1 & \text{otherwise} \end{cases}$$

حيث أن  $\varepsilon$  مقدار موجب صغير، إذن تكون دالة المخاطرة كالاتي:

$$\mathfrak{R}(\hat{\theta} | y_1, y_2, \dots, y_n) = \int_{\theta \in \Omega} P(\theta | y_1, y_2, \dots, y_n) d\theta$$

وعند التعويض بدالة الخسارة بما يساويها نحصل على الآتي:

$$\begin{aligned}\mathfrak{R}(\hat{\theta} | y_1, y_2, \dots, y_n) &= \int_{|\theta - \hat{\theta}| > \varepsilon} P(\theta | y_1, y_2, \dots, y_n) d\theta \\ &= 1 - \int_{\hat{\theta} - \varepsilon}^{\hat{\theta} + \varepsilon} P(\theta | y_1, y_2, \dots, y_n) d\theta\end{aligned}$$

وبتطبيق مبرهنة القيمة المتوسطة (Mean Value Theorem) - الملحق A - على التكامل الأخير نحصل على. [Melsa & Cohn, 1978]:

$$\begin{aligned}\mathfrak{R}(\hat{\theta} | y_1, y_2, \dots, y_n) &= 1 - [(\hat{\theta} + \varepsilon) - (\hat{\theta} - \varepsilon)] \times P(\hat{\theta} | y_1, y_2, \dots, y_n) \\ &= 1 - 2\varepsilon \times P(\hat{\theta} | y_1, y_2, \dots, y_n)\end{aligned}$$

ولأن مقدر بيز بشكل عام يجعل دالة المخاطرة أقل ما يمكن- أنظر الشكل (٢-١) -، إذن ذلك يتطلب أن يكون التوزيع اللاحق أكبر ما يمكن، أي أن:

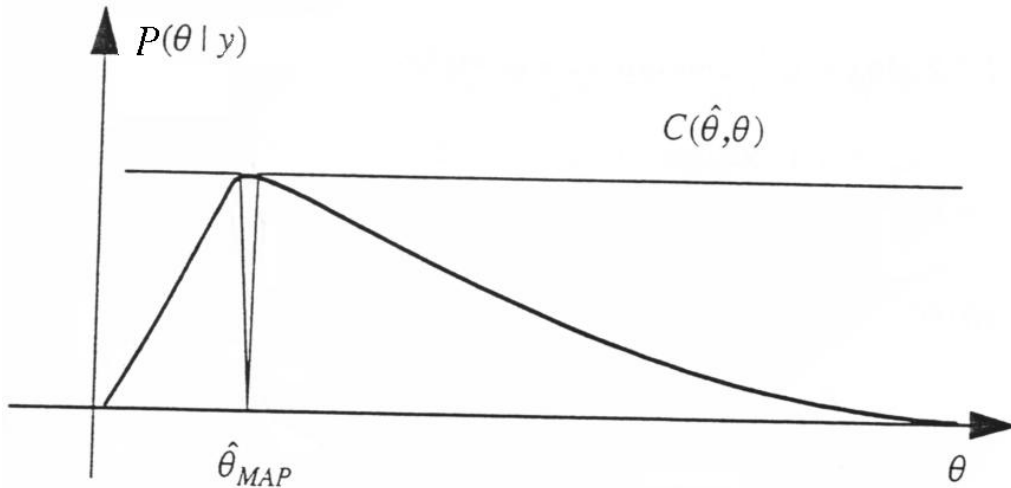
$$P(\hat{\theta}_B | y_1, y_2, \dots, y_n) \geq P(\hat{\theta} | y_1, y_2, \dots, y_n), \forall (\hat{\theta} \in \Omega)$$

حيث أن  $\hat{\theta} \neq \hat{\theta}_B$ ، إذن يكون مقدر بيز لتعظيم التوزيع اللاحق:

$$\hat{\theta}_{MAP} = \arg \max P(\theta | y_1, y_2, \dots, y_n)$$

معنى ذلك أن مقدر بيز لتعظيم التوزيع اللاحق هو منوال التوزيع اللاحق

.Mode of Posterior Distribution



الشكل (٢-١)

دالة الكلفة المنتظمة ومنحنى التوزيع اللاحق. [Vaseghi, 2000]

### ٢-٥-١- تقدير الإمكان الأعظم (Maximum Likelihood Estimation):

وهو التقدير الذي يجعل من دالة الإمكان أعظم ما يمكن ويرمز له بـ  $\hat{\theta}_{ML}$ ، وهذا المقدر لا يمكن إيجاده في استدلال بيز إلا في حالة كون دالة كلفة الخطأ منتظمة وكذلك يجب أن يكون التوزيع الأولي منتظماً أيضاً أو قليل المعلومات، فعندئذ يكون التوزيع اللاحق في تناسب مع دالة الإمكان، فنتمكن من إيجاد مقدر الإمكان الأعظم وذلك بالتعويض عن دالة الخسارة بما يساويها وعن التوزيع اللاحق بما يساويه من مبرهنة بيز ومن ثم تطبيق مبرهنة القيمة المتوسطة على التكامل:

$$\begin{aligned}
\mathfrak{R}(\hat{\theta} | y_1, y_2, \dots, y_n) &= \int_{\theta \in \Omega} [C(\hat{\theta}, \theta)] P(\theta | y_1, y_2, \dots, y_n) d\theta \\
&= \text{Const.} \times \int_{|\theta - \hat{\theta}| > \varepsilon} P(\theta) L(\theta) d\theta \\
&= \text{Const.} \times [1 - 2\varepsilon \times L(\hat{\theta})]
\end{aligned}$$

من الواضح أن جعل دالة المخاطرة أقل ما يمكن يتطلب جعل دالة الإمكان أكبر ما يمكن:

$$\hat{\theta}_{ML} = \arg \max L(\theta)$$

إن الفرق الأساسي بين مقدر الإمكان الأعظم ومقدر بيز لتعظيم التوزيع اللاحق يكمن في أن مقدر الإمكان الأعظم يفترض أن يكون التوزيع الأولي منتظماً- قيمة ثابتة على جميع قيم  $\theta$  - أو قليل المعلومات. وجدير بالذكر أن مقدر الإمكان الأعظم في هذه الحالة هو نفسه مقدر الإمكان الأعظم في الاستدلال الكلاسيكي. [Vaseghi, 2000]

### ١-٥-٣- تقدير بيز ذو أقل متوسط مربع خطأ:

#### Minimum Mean Square Error Bayes Estimation:

يكون مقدر بيز في هذه الطريقة هو المقدر الذي يجعل توقع دالة كلفة مربع الخطأ أو دالة الخسارة التربيعية Quadratic Loss Function أقل ما يمكن، ويرمز له بـ  $\hat{\theta}_{MMSE}$ . حيث أن دالة كلفة مربع الخطأ هي:

$$C(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$$

إذن:

$$\begin{aligned}
\mathfrak{R}(\hat{\theta} | y_1, y_2, \dots, y_n) &= E[(\hat{\theta} - \theta)^2 | y_1, y_2, \dots, y_n] \\
&= \int_{\theta \in \Omega} (\hat{\theta} - \theta)^2 P(\theta | y_1, y_2, \dots, y_n) d\theta
\end{aligned}$$

ولإيجاد مقدر بيز الذي له أقل متوسط مربع خطأ (أو الأمثل Optimal) نشق دالة المخاطرة الشرطية بالنسبة لـ  $\hat{\theta}$  ونساوي المقدار بالصفر وكالاتي:

$$\begin{aligned}
\frac{\partial \mathfrak{R}(\hat{\theta} | y_1, y_2, \dots, y_n)}{\partial \hat{\theta}} &= 2\hat{\theta} \int_{\theta \in \Omega} P(\theta | y_1, y_2, \dots, y_n) d\theta - 2 \int_{\theta \in \Omega} \theta P(\theta | y_1, y_2, \dots, y_n) d\theta \\
\therefore \hat{\theta}_{MMSE} &= \int_{\theta \in \Omega} \theta P(\theta | y_1, y_2, \dots, y_n) d\theta = E(\theta | y_1, y_2, \dots, y_n)
\end{aligned}$$

وكما نلاحظ فإن تقدير بيز ذو أقل متوسط مربع خطأ هو نفسه معدل التوزيع اللاحق، ويعد هذا المقدر من أكثر مقدرات بيز استخداماً وذلك لسهولة إيجاده.

### ١-٥-٤- تقدير بيز ذو أقل قيمة متوسط خطأ مطلق:

#### Minimum Mean Absolute Value of Error Bayes Estimation:

وهو مقدر بيز الذي يجعل توقع دالة كلفة الخطأ المطلق (دالة المخاطرة) أقل ما يمكن، ويرمز

له بـ  $\hat{\theta}_{MAVE}$ ، وتعرف دالة كلفة الخطأ المطلق كالآتي:

$$C(\hat{\theta}, \theta) = |\hat{\theta} - \theta|$$

$$\begin{aligned} \mathfrak{R}(\hat{\theta} | y_1, y_2, \dots, y_n) &= \int_{\theta \in \Omega} |\hat{\theta} - \theta| P(\theta | y_1, y_2, \dots, y_n) d\theta \\ &= \int_{-\infty}^{\hat{\theta}} (\hat{\theta} - \theta) P(\theta | y_1, y_2, \dots, y_n) d\theta + \int_{\hat{\theta}}^{\infty} (\theta - \hat{\theta}) P(\theta | y_1, y_2, \dots, y_n) d\theta \end{aligned}$$

وبأخذ المشتقة لدالة مخاطرة بيز بالنسبة لـ  $\hat{\theta}$  ومساواة الناتج بالصفر نحصل على الآتي:

$$\frac{\partial \mathfrak{R}(\hat{\theta} | y_1, y_2, \dots, y_n)}{\partial \hat{\theta}} = \int_{-\infty}^{\hat{\theta}} P(\theta | y_1, y_2, \dots, y_n) d\theta - \int_{\hat{\theta}}^{\infty} P(\theta | y_1, y_2, \dots, y_n) d\theta$$

$$\int_{-\infty}^{\hat{\theta}} P(\theta | y_1, y_2, \dots, y_n) d\theta = \int_{\hat{\theta}}^{\infty} P(\theta | y_1, y_2, \dots, y_n) d\theta$$

إذن تكون قيمة  $\hat{\theta}_{MAVE}$  هي القيمة التي تحقق المعادلة الأخيرة، وبما أن الوسيط هو الذي يحقق

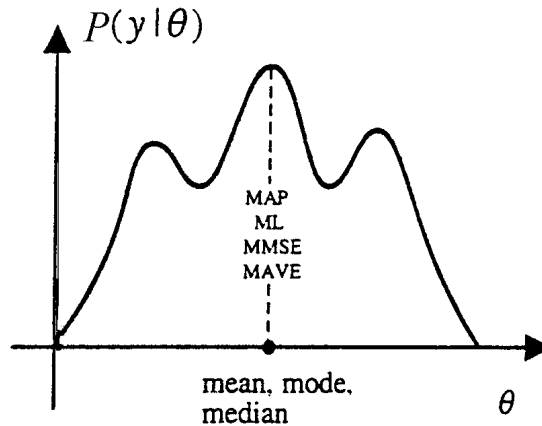
المعادلة أعلاه، فيكون مقدر بيز في هذه الحالة هو وسيط التوزيع اللاحق.

■ ملاحظة:

عندما يكون التوزيع متمائل (Symmetric) ويكون التوزيع الأولي منتظماً أو قليل المعلومات فإن

المقدرات الأربعة تتساوى -أنظر الشكل (٣-١)-، أي أن:  $\hat{\theta}_{MAP} = \hat{\theta}_{ML} = \hat{\theta}_{MMSE} = \hat{\theta}_{MAVE}$

وللمزيد من المعلومات أنظر [Vaseghi, 2000 ; Melsa & Cohn, 1978].



الشكل (٣-١)

منحنى التوزيع المتمائل إضافة إلى الوسط والوسيط والمنوال. [Vaseghi, 2000]

## ٦-١- تحليل بيز المتسلسل (Bayesian Sequential Analysis):

يقصد بالتحليل المتسلسل أخذ المشاهدات بشكل متسلسل - واحدة تلو الأخرى - بدون تحديد مسبق لحجم العينة، والغاية من ذلك الحصول على دقة أعلى مقابل كلفة أدنى، لذا فإن هذا الأسلوب كثير الاستخدام في مجال السيطرة النوعية. وكان [Wald, 1947] أول من استخدم التحليل المتسلسل في مجال اختبار الفرضيات.

ولغرض توضيح التحليل المتسلسل بشكل مبسط نفترض المثال الآتي:

لنفرض أنه لدينا المشاهدة  $y_1$  من مجتمع له معلمة  $\theta$  لها توزيع أولي  $P(\theta)$ ، وقمنا بحساب التوزيع اللاحق للمعلمة، فعندئذ:

$$P(\theta | y_1) = P(\theta)P(y_1 | \theta)$$

ولنفرض أيضاً أننا أخذنا مشاهدة أخرى  $y_2$  من نفس مجتمع  $y_1$  ومستقلة عنها، وأردنا حساب

التوزيع اللاحق للمعلمة، فعندئذ سوف نستخدم  $P(\theta | y_1)$  كتوزيع أولي، وكالآتي:

$$P(\theta | y_1, y_2) \propto P(\theta | y_1)P(y_2 | \theta)$$

ومن هنا يتضح المفهوم الأساسي لتحليل بيز المتسلسل، حيث يلعب التوزيع اللاحق في المرحلة

$m$  دور التوزيع الأولي في المرحلة  $m+1$ ، والآن إذا كان لدينا  $(n)$  من المشاهدات المستقلة أخذت بشكل متسلسل فإن التوزيع اللاحق في كل مرحلة يحسب كما يأتي:

$$P(\theta | y_1, \dots, y_m) \propto P(y_m | \theta)P(\theta | y_1, \dots, y_{m-1}), m = 2, \dots, n$$

حيث أن:

$$P(\theta | y_1) \propto P(y_1 | \theta)P(\theta)$$

### ■ ملاحظة:

إن استخدام توزيع أولي جديد للمعلمة في كل مرحلة لا يعني أن المعلومات الابتدائية قد تلاشت بل تكون مضمنة في كل مرحلة في التوزيع الأولي الخاص بتلك المرحلة، ويلاحظ أيضاً أن التوزيع الأولي في التحليل المتسلسل يكون دائماً من نوع Informative Prior. أنظر [Wald, 1947]

## ٧-١- النماذج الخطية الحركية (Dynamic Linear Models):

إن النماذج الخطية الحركية هي منظومة من المعادلات توضح العلاقة التصادفية بين المتغيرات العشوائية والمعلمت أو متغير الحالة (State Variable) في عمليات السلاسل الزمنية، فضلاً عن ذلك توضح معلمت العملية والتطورات التي تطرأ عليها نتيجة مرور الزمن ويمكن التعبير عن ذلك رياضياً كما في منظومة المعادلات (١،٤) أدناه. [Harrison & Stevens, 1976]



وهناك الكثير من الأمثلة عن السلاسل الزمنية في الحقول المختلفة مثل الاقتصاد وهندسة الاتصالات والعلوم الطبيعية والعسكرية وغيرها الكثير. [Dulger, 2001]

حيث أن النمذجة الرياضية والإحصائية لهذه العمليات تعتمد على أصناف متنوعة من النماذج الحركية، وقد تمّ استعراض العديد من هذه النماذج من قبل الباحثين (Harrison & Stevens) وبأسلوب سلس، ومنها نماذج الانحدار الساكنة (Static Regression Models) وكذلك نماذج الانحدار الحركية (Dynamic Regression Models) والنماذج الموسمية (Seasonal Models) والنماذج المستقرة (Steady Models) وغيرها من النماذج الحركية.

إن أكثر أصناف النماذج الخطية الحركية استخداماً هي النماذج الخطية الحركية الطبيعية (Normal Dynamic Linear Models) ويشار إليها اختصاراً بـ (DLMs) عندما تكون الحالة الطبيعية مدركة. [West & Harrison, 1989]

ويتألف النموذج الخطي الحركي الطبيعي بصورة عامة من معادلتين خطيتين، تدعى الأولى بمعادلة المشاهدة (Observation Equation) والثانية بمعادلة النظام (System Equation)، وكالاتي:

$$Y_t = F_t \theta_t + v_t \quad \dots\dots\dots(1.4a)$$

$$\theta_t = G_t \theta_{t-1} + \omega_t \quad \dots\dots\dots(1.4b)$$

حيث أن:

$t$  : يمثل دليل الزمن، ويأخذ القيم ( $t = 1, 2, \dots$ ).

$Y_t$  : يمثل متجه المشاهدات عند الزمن  $t$ ، وهو ذو بعد  $(m \times 1)$ .

$\theta_t$  : يمثل متجه المعلمات عند الزمن  $t$ ، ويعرف بمتجه النظام أو الحالة، وهو ذو بعد  $(n \times 1)$ .

$F_t$  : تمثل مصفوفة المتغيرات المستقلة، وتكون معلومة عند الزمن  $t$ ، وهي ذات بعد  $(m \times n)$ .

$G_t$  : تمثل مصفوفة النظام، وتكون معلومة عند الزمن  $t$ ، وهي ذات بعد  $(n \times n)$ .

$v_t$  : يمثل متجه الأخطاء العشوائية الطبيعية للمشاهدات بمتوسط  $E(v_t) = 0$  ومصفوفة

تباين  $V_t = E(v_t v_t')$  معلومة عند الزمن  $t$ ، أي أن  $v_t \sim N_m(0, V_t)$ ، وهو ذو بعد  $(m \times 1)$ .

$\omega_t$  : يمثل متجه الأخطاء العشوائية الطبيعية للنظام بمتوسط  $E(\omega_t) = 0$  ومصفوفة

تباين  $W_t = E(\omega_t \omega_t')$  معلومة عند الزمن  $t$ ، أي أن  $\omega_t \sim N_n(0, W_t)$ ، وذو بعد  $(n \times 1)$ .

ويمكن وصف النموذج الخطي الحركي الطبيعي بواسطة الرباعي (Quadruple) الآتي:  
 $\{F_t, G_t, V_t, W_t\}$

حيث أن  $V_t$  و  $W_t$  هما مصفوفتا التباين لخطأ المشاهدة وخطأ النظام وهما ذاتا بعد  $(m \times m)$  و  $(n \times n)$  على التوالي.

#### ■ ملاحظة:

١. إذا كان كل من المصفوفتين  $F_t$  و  $G_t$  ثابتتين لجميع قيم  $t$ ، فعندئذ يدعى النموذج بنموذج السلسلة الزمنية الخطية الحركي (Time Series DLM or TSDLM)، ويمكن وصفه من خلال الرباعي  $\{F, G, V_t, W_t\}$ .

٢. إذا كان كل من المصفوفتين  $V_t$  و  $W_t$  ثابتتين لجميع قيم  $t$  في نموذج السلسلة الزمنية الخطية الحركي، فعندئذ يدعى النموذج بالنموذج الخطية الحركي الثابت Constant DLM والذي يمكن وصفه من خلال الرباعي  $\{F, G, V, W\}$ .

إن هذه التصنيفات المهمة تضم الجوهر الأساسي لجميع نماذج السلاسل الزمنية الخطية الكلاسيكية. [West & Harrison, 1989]

#### ❖ بعض فوائد النماذج الخطية الحركية:

١. تفادي فرضية الاستقرارية (Stationary) عند بناء النماذج، ففي كثير من التطبيقات توفر النمذجة بأسلوب حركي (Dynamically) مرونة عالية في تمثيل السلاسل الزمنية.

٢. تمكننا من نمذجة المكونات المختلفة للسلسلة الزمنية (مثل نماذج الانحدار والنماذج الموسمية...) بشكل منفصل بنماذج جزئية (Sub-Models) ومن ثم توضع هذه النماذج معاً لصياغة النموذج العام للسلسلة الزمنية.

٣. وكذلك في استخدام التقنيات التحليلية للنماذج الخطية الحركية مثل مرشح كالمن (Kalman Filter) الذي يزودنا بالمقدر الأمثل (وبنفس الوقت يزودنا بالتوقع الشرطي لمتجه الحالة) تحت الفروض التقليدية، أو يمكن أن يعد حلاً تقريبياً في الحالات الأكثر عمومية.

انظر [Campagnoli, et al. 2000]

#### ١-٨- مبرهنة بيز والنماذج الخطية الحركية:

إن الاستدلال حول قيمة  $(\theta_t)$  عند الزمن  $(t)$  في النماذج الخطية الحركية يتطلب توافر معلومات مسبقة من المرحلة الزمنية التي تسبقها  $(t-1)$ ، وهذه المعلومات يمكن تمثيلها بالتوزيع اللاحق في المرحلة  $(t-1)$ ، كما يتطلب معرفة توزيع كل من حدي الخطأ  $(\nu_t \& \omega_t)$ ، وبالتأكيد يتطلب أيضاً

مشاهدة جديدة ( $Y_t$ ). [Harrison & Stevens, 1976]. وبذلك يمكننا تطبيق أسلوب بيز في التحليل المتسلسل، حيث يتم حساب التوزيعين الأولي واللاحق لـ ( $\theta_t$ ) في كل مرحلة زمنية ( $t$ ). والفقرة الآتية توضح العمليات الرياضية المتبعة لإيجاد التوزيعين المذكورين في كل مرحلة زمنية ( $t$ ).

#### ❖ إيجاد التوزيعين الأولي واللاحق للنظام عند الزمن $t$ :

إن النموذج الخطي الحركي الطبيعي هو النموذج الخطي الحركي الذي تتوزع مكوناته وفق التوزيع الطبيعي، أي أن المشاهدات والمعلمات فضلاً عن الأخطاء تتبع التوزيع الطبيعي عند كل زمن. فإذا فرضنا أن التمثيل الاحتمالي للمعلومات المتوافرة حول المعلمة  $\theta_t$  عند الزمن  $t=0$  يكون كالآتي:

$$\theta_0 \sim N_n(M_0, C_0)$$

حيث أن:

$$\theta_0 : \text{يمثل متجه المعلمات عند الزمن } t=0, \text{ ذو بعد } (n \times 1).$$

$$M_0 : \text{يمثل متجه المتوسط عند الزمن } t=0, \text{ ذو بعد } (n \times 1).$$

$$C_0 : \text{تمثل مصفوفة التباين عند الزمن } t=0, \text{ ذو بعد } (n \times n).$$

ولنفرض أن ( $D_{t-1}$ ) تشمل جميع المعلومات المتأتية من المشاهدات ( $Y_1, Y_2, \dots, Y_{t-1}$ ) ومن مصفوفتي التباين  $V_t$  و  $W_t$  لجميع قيم  $t$  فضلاً عن المعلومات المتوافرة عند الزمن ( $t=0$ ) أي  $D_0$ ، أي أن:

$$D_{t-1} = \{Y_1, Y_2, \dots, Y_{t-1}, V_t, W_t, D_0\}$$

لذا فإن المعلومات الجديدة عند أي زمن  $t$  تكون متأتية من المشاهدة الجديدة  $Y_t$  فضلاً عن المعلومات الأولية ( $D_{t-1}$ ). [West & Harrison, 1989]. أي أن:

$$D_t = \{Y_t, D_{t-1}\}$$

وباستخدام الاستقراء الرياضي فإن التوزيع اللاحق للنظام عند الزمن  $t-1$  هو:

$$(\theta_{t-1} | D_{t-1}) \sim N_n(M_{t-1}, C_{t-1})$$

ومن مبرهنة بيز باستخدام التحليل المتسلسل، فإن التوزيع الاحتمالي اللاحق للنظام عند الزمن  $t$  يمكن إيجاده حسب الصيغة الآتية:

$$P(\theta_t | D_t) \propto P(\theta_t | D_{t-1})P(Y_t | \theta_t) \dots \dots \dots (1.5)$$

من الواضح أن  $(\theta_t | D_{t-1})$  يتبع التوزيع الطبيعي وذلك لأن  $\theta_t$  دالة خطية لـ  $(\omega_t$  و  $\theta_{t-1})$  وهما يتوزعان توزيعاً طبيعياً مستقلاً، ويمكن إيجاد معاملات التوزيع الأولي كالاتي:

$$\begin{aligned} E(\theta_t | D_{t-1}) &= E(G_t \theta_{t-1} + \omega_t | D_{t-1}) \\ &= G_t M_{t-1} \\ \text{Var}(\theta_t | D_{t-1}) &= \text{Var}(G_t \theta_{t-1} + \omega_t | D_{t-1}) \\ &= G_t C_{t-1} G_t' + W_t \end{aligned}$$

ويرمز لهذا التباين بـ  $(R_t)$ ، وكالاتي:

$$R_t = G_t C_{t-1} G_t' + W_t \quad \dots\dots\dots(1.6)$$

إذن يكون التوزيع الأولي للنظام عند الزمن  $(t)$  توزيعاً طبيعياً بمتجه متوسط  $(G_t M_{t-1})$  وهو ذو بعد  $(n \times 1)$ ، ومصفوفة تباين  $(R_t)$  ذو بعد  $(n \times n)$ ، أي أن:

$$\therefore (\theta_t | D_{t-1}) \sim N_n(G_t M_{t-1}, R_t) \quad \dots\dots\dots(1.7)$$

وكذلك بما أن  $(Y_t | \theta_t)$  هو عبارة عن دالة خطية لـ  $(\nu_t$  و  $\theta_t)$  وهما يتوزعان توزيعاً طبيعياً مستقلاً، إذن يكون توزيع  $(Y_t | \theta_t)$  طبيعياً أيضاً، ويتم إيجاد معاملات التوزيع كالاتي:

$$\begin{aligned} E(Y_t | \theta_t) &= E(F_t \theta_t + \nu_t | \theta_t) \\ &= F_t \theta_t \\ \text{Var}(Y_t | \theta_t) &= \text{Var}(F_t \theta_t + \nu_t | \theta_t) \\ &= V_t \end{aligned}$$

$$\therefore (Y_t | \theta_t) \sim N_m(F_t \theta_t, V_t) \quad \dots\dots\dots(1.8)$$

وبالعودة إلى أسلوب بيز في إيجاد التوزيع اللاحق في العلاقة (١،٥) وباستخدام (١،٧) و(١،٨)، نحصل على:

$$\begin{aligned} P(\theta_t | D_t) &\propto \exp - \frac{1}{2} [(\theta_t - G_t M_{t-1})' R_t^{-1} (\theta_t - G_t M_{t-1}) + (Y_t - F_t \theta_t)' V_t^{-1} (Y_t - F_t \theta_t)] \\ &= \exp - \frac{1}{2} [\theta_t' R_t^{-1} \theta_t - 2 M_{t-1}' G_t' R_t^{-1} \theta_t + M_{t-1}' G_t' R_t^{-1} G_t M_{t-1} \\ &\quad + Y_t V_t^{-1} Y_t - 2 Y_t V_t^{-1} F_t \theta_t + \theta_t' F_t V_t^{-1} F_t \theta_t] \\ &= \exp - \frac{1}{2} [\theta_t' (R_t^{-1} + F_t V_t^{-1} F_t) \theta_t - 2 (M_{t-1}' G_t' R_t^{-1} + Y_t V_t^{-1} F_t) \theta_t + Const.] \end{aligned}$$

حيث أن  $Const.$  لا يحتوي على  $\theta_t$ .

وبإكمال المربع وحذف الحدود التي لا تحتوي على  $\theta_t$  نتوصل إلى:

$$P(\theta_t | D_t) \propto \exp - \frac{1}{2} [(\theta_t - M_t)' C_t^{-1} (\theta_t - M_t)]$$

حيث أن:

$$M_t = G_t M_{t-1} + R_t F_t' (F_t R_t F_t' + V_t)^{-1} (Y_t - F_t G_t M_{t-1}) \quad \dots\dots\dots(1.9)$$

$$C_t = (R_t^{-1} + F_t V_t^{-1} F_t)^{-1} \quad \dots\dots\dots(1.10)$$

إذن يكون التوزيع اللاحق للنظام عند الزمن  $t$  توزيعاً طبيعياً متعدداً بمتجه متوسط  $M_t$  ذو بعد  $(n \times 1)$ ، ومصفوفة تباين  $C_t$  ذو بعد  $(n \times n)$ ، ويمكن تمثيله وصفاً كالآتي:

$$\therefore (\theta_t | D_t) \sim N_n(M_t, C_t) \quad \dots\dots\dots(1.11)$$

أنظر [West & Harrison, 1989; Jalil, 1988]

إن طريقة التقدير هذه تعرف بالتقدير التعاقبي (Recursive Estimation) وتدعى أحياناً بنظرية التقدير الحديثة (Modern Estimation Theory). [Bozic, 1979].

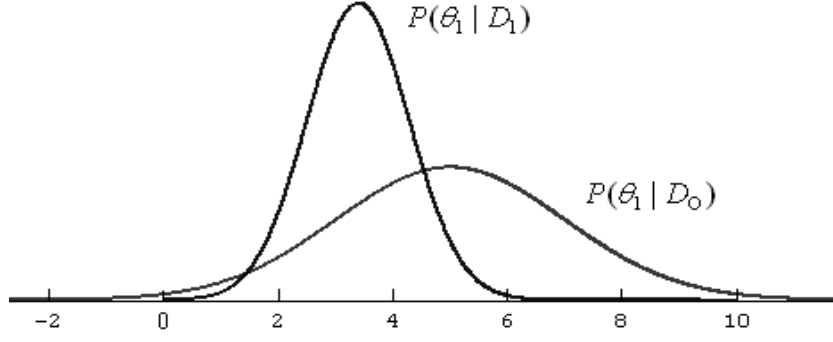
#### ■ ملاحظة:

قد يكون من المفيد استخدام متطابقة مأخوذة معكوس المصفوفة (Matrix Inversion Lemma) وذلك لصياغة مصفوفة تباين التوزيع اللاحق  $C_t$  بصيغة أخرى تكون في أغلب الأحيان أكثر ملائمة حسابياً، [Melsa & Cohn, 1978]، وكالآتي:

$$(R_t^{-1} + F_t V_t^{-1} F_t)^{-1} = R_t - R_t F_t' (F_t R_t F_t' + V_t)^{-1} F_t R_t = C_t \quad \dots\dots\dots(1.12)$$

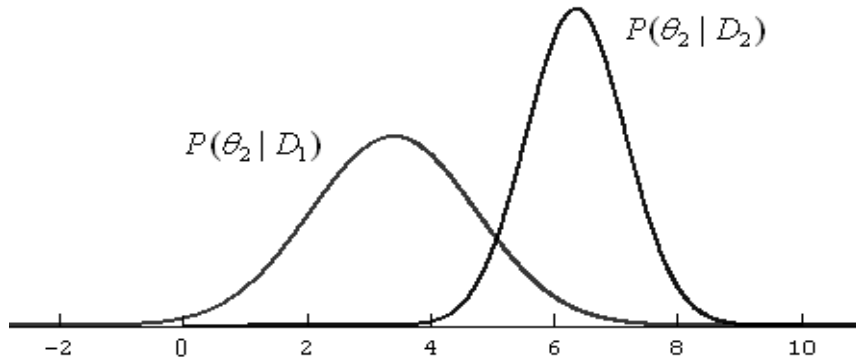
#### ◀ توضيح:

افرض أنه لدينا النموذج الخطي الحركي الموصوف بالرباعي  $\{1,1,1,1\}$ ، وافرض أيضاً أن  $(t=1,2)$  وأن  $(y_1=3)$  و  $(y_2=8)$ ، وليكن التوزيع الابتدائي (Initial Distribution) للمعلمة هو  $\theta_0 \sim N(5,3)$ . فعند الزمن  $(t=1)$  يكون التوزيع الأولي للمعلمة هو  $(\theta_1 | D_0) \sim N(5,4)$ ، حيث يمكن إيجاده بسهولة من (1.7)، وبالإمكان إيجاد التوزيع اللاحق للمعلمة عند الزمن  $(t=1)$  وذلك باستخدام  $(1,1)$  ويكون  $(\theta_1 | D_1) \sim N(3.4,0.8)$ ، ويلاحظ أن متوسط التوزيع اللاحق أقل من متوسط التوزيع الأولي وذلك لأن قيمة  $y_1$  أقل من متوسط التوزيع الأولي، لا حظ الشكل (٤-١).



الشكل (٤-١)  
منحنى التوزيعين الأولي واللاحق عند الزمن  $(t = 1)$

وفي الزمن  $(t=2)$  يمكن إيجاد التوزيع الأولي للمعلمة من  $(1.7)$  وهو  $(\theta_2 | D_1) \sim N(3.4, 1.8)$ ، وكذلك يمكن إيجاد التوزيع اللاحق للمعلمة من  $(1, 1)$  ويكون  $(\theta_2 | D_2) \sim N(6.357, 0.642)$ ، وبما أن  $(y_2 = 8)$  فإن منحنى التوزيع اللاحق ينسحب نحو اليمين مقارنة بمنحنى التوزيع الأولي، لاحظ الشكل (٥-١).



الشكل (٥-١)  
منحنى التوزيعين الأولي واللاحق عند الزمن  $(t = 2)$

### ٩-١- بعض النماذج الخطية الحركية:

إن التطبيقات العملية الإحصائية مثل السيطرة النوعية (Quality Control)، معالجة الصورة (Image Processing) ومعالجة الإشارة (Signal Processing) تتبع أنواعاً مختلفة من النماذج الحركية، وذلك تبعاً للتغيرات الحاصلة خلال الزمن في متجه الحالة بشكل خاص، وبالعملية العشوائية بشكل عام. ومن هذه النماذج:

### ١-٩-١ - النموذج الدرّجي (Step Model):

وهو أحد أبسط النماذج الخطية الحركية وأكثرها انتشاراً في كثير من التطبيقات، وفيه تحدث تغيرات عشوائية مفاجئة في مستوى العملية يطلق عليها قفزات (Jumps). وفي كل قفزة يتغير مستوى العملية إلى مستوى آخر جديد وذلك بإضافة حد الخطأ العشوائي الطبيعي  $\delta_{\mu_t}$ .

[Jalil, 1988]

وهذا لا يعني وجود تغير في مستوى العملية عند كل فترة زمنية  $(t-1, t)$  فقد يبقى المستوى ثابتاً لفترة زمنية واحدة أو أكثر ومن ثمّ تحدث القفزة بشكل مفاجئ، إن هذا التغير يظهر في تباين  $\mu_t$  حيث يصبح التباين أكبر، وتكون صيغة هذا النموذج كالآتي:

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t & \varepsilon_t &\overset{i.i.d.}{\sim} N(0, \sigma_\varepsilon^2) \\ \mu_t &= \mu_{t-1} + \delta_{\mu_t} & \delta_{\mu_t} &\overset{i.i.d.}{\sim} N(0, \sigma_{\delta_\mu}^2) \end{aligned}$$

حيث أن  $y_t$  تمثل القيمة المشاهدة عند الزمن  $t$ ، وتمثل  $\mu_t$  مستوى (معدل) العملية الحقيقي عند الزمن  $t$ ، ويمثل كلٌّ من  $(\varepsilon_t$  و  $\delta_{\mu_t})$  حد الخطأ العشوائي لمعادلة المشاهدة ومعادلة النظام على التوالي وهما يتوزعان توزيعاً طبيعياً بمتوسط (0) وتباين  $(\sigma_\varepsilon^2)$  و  $(\sigma_{\delta_\mu}^2)$  على التوالي ومستقلين عن بقية المتغيرات فضلاً عن عدم ترابطهما معاً. ويلاحظ في هذا النموذج أن  $(n=m=1)$ . ويمكن وصف هذا النموذج من خلال الرباعي  $\{1, 1, \sigma_\varepsilon^2, \sigma_{\delta_\mu}^2\}$ . حيث أن:

$$(F_t = F = 1), (G_t = G = 1), (V_t = \sigma_\varepsilon^2), (W_t = \sigma_{\delta_\mu}^2)$$

❖ إيجاد التوزيعين الأولي واللاحق لمستوى العملية  $\mu_t$  عند الزمن  $t$ :

مما تقدّم فإن مستوى العملية يتبع التوزيع الطبيعي بمتوسط  $(M_0 = m_0)$  وتباين  $(C_0 = c_0)$ ، أي أن:

$$\mu_0 \sim N(m_0, c_0)$$

وباستخدام الاستقراء الرياضي، فإن:

$$(\mu_{t-1} | D_{t-1}) \sim N(m_{t-1}, c_{t-1})$$

ومن (1.7) نحصل على متوسط وتباين التوزيع الأولي في الزمن  $t$ ، وكالآتي:

$$E(\theta_t | D_{t-1}) = m_{t-1}$$

$$V(\theta_t | D_{t-1}) = R_t = r_t = c_{t-1} + \sigma_{\delta\mu}^2$$

ومن (1.11) نحصل على متوسط وتباين التوزيع اللاحق في الزمن  $t$ ، وكالاتي:

$$M_t = m_t = \frac{(y_t/\sigma_\varepsilon^2) + (m_{t-1}/r_t)}{(1/\sigma_\varepsilon^2) + (1/r_t)}$$

$$C_t = c_t = \frac{r_t \sigma_\varepsilon^2}{r_t + \sigma_\varepsilon^2}$$

أي أن:

$$(\mu_t | D_t) \sim N(m_t, c_t)$$

### ٢-٩-١- نموذج النمو الخطي الحركي المركب:

#### Combined Dynamic Linear Growth Model:

يعد هذا النموذج نموذجاً موسعاً للنموذج الدرجي، حيث يضاف إليه حد النمو (عامل الميل)  $\beta_t$  الذي يتغير بتغير الزمن من خلال حد الخطأ العشوائي  $\delta_{\beta_t}$ . [Harrison & Stevens, 1976]. فضلاً عن التغيرات الفجائية التي تحصل في مستوى (معدل) العملية  $\mu_t$  خلال الزمن، يحدث نمو تزايدي (Incremental Growth) في مستوى العملية خلال الزمن أي التغيير التدريجي (Gradual). وهناك العديد من التطبيقات في الاقتصاد (Economics) وسيطرة العملية (Process Control) التي تتوافق مع هذا النموذج. ويمكن كتابة النموذج كالاتي:

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t & \varepsilon_t &\stackrel{i.i.d.}{\sim} N(0, \sigma_\varepsilon^2) \\ \mu_t &= \mu_{t-1} + \beta_t + \delta_{\mu_t} & \delta_{\mu_t} &\stackrel{i.i.d.}{\sim} N(0, \sigma_{\delta\mu}^2) \\ \beta_t &= \beta_{t-1} + \delta_{\beta_t} & \delta_{\beta_t} &\stackrel{i.i.d.}{\sim} N(0, \sigma_{\delta\beta}^2) \end{aligned}$$

حيث أن  $y_t$  تمثل القيمة المشاهدة عند الزمن  $t$ ، وتمثل  $\mu_t$  مستوى (معدل) العملية عند الزمن  $t$ ، وتمثل  $\beta_t$  عامل ميل المستوى عند الزمن  $t$ ، ويمثل كلٌّ من  $\varepsilon_t$  و  $\delta_{\mu_t}$  و  $\delta_{\beta_t}$  حد الخطأ العشوائي الطبيعي للمشاهدة وللمستوى العملية ولميل مستوى العملية على التوالي، وتكون الأخطاء مستقلة عن بعضها البعض. ويلاحظ في هذا النموذج أن:

$$\begin{aligned} Y_t = y_t, \theta_t &= \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix}, F_t = F = \begin{bmatrix} 1 & 0 \end{bmatrix}, G_t = G = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}; \\ V_t = \sigma_\varepsilon^2, W_t &= \begin{bmatrix} \sigma_{\delta\mu}^2 + \sigma_{\delta\beta}^2 & \sigma_{\delta\beta}^2 \\ \sigma_{\delta\beta}^2 & \sigma_{\delta\beta}^2 \end{bmatrix}, v_t = \varepsilon_t, \omega_t = \begin{bmatrix} \delta_{\mu_t} + \delta_{\beta_t} \\ \delta_{\beta_t} \end{bmatrix}, m=1, n=2 \end{aligned}$$



ويمكن كتابة هذا النموذج بطريقة المصفوفات وكالاتي:

$$\begin{aligned} Y_t &= F\theta_t + v_t \\ \theta_t &= G\theta_{t-1} + \omega_t \end{aligned} \quad \dots\dots\dots(1.13)$$

❖ إيجاد التوزيعين الأولي واللاحق لمتجه الحالة  $\theta_t$  عند الزمن  $t$ :

مما تقدم فإن النظام يتبع التوزيع الطبيعي الثنائي المتغير الآتي:

$$\theta_0 \sim N_2(M_0, C_0)$$

ومن (١,٦) فإن مصفوفة تباين التوزيع الأولي  $R_t$  يمكن كتابتها كالاتي:

$$\begin{aligned} R_t &= \begin{bmatrix} r_{11(t)} & r_{12(t)} \\ r_{21(t)} & r_{22(t)} \end{bmatrix} \\ &= \begin{bmatrix} c_{11(t-1)} + 2c_{12(t-1)} + c_{22(t-1)} + \sigma_{\delta_\mu}^2 + \sigma_{\delta_\beta}^2 & c_{12(t-1)} + c_{22(t-1)} + \sigma_{\delta_\beta}^2 \\ c_{12(t-1)} + c_{22(t-1)} + \sigma_{\delta_\beta}^2 & c_{22(t-1)} + \sigma_{\delta_\beta}^2 \end{bmatrix} \dots\dots(1.14) \end{aligned}$$

ومن (1.11) وبالاعتماد على الصيغة (1.12) نحصل على متجه المتوسط ومصفوفة التباين للتوزيع اللاحق في الزمن  $t$ ، وكالاتي:

$$\begin{aligned} M_t &= \begin{bmatrix} m_t \\ b_t \end{bmatrix} = \begin{bmatrix} m_{t-1} + b_{t-1} + \frac{r_{11(t)}}{r_{11(t)} + \sigma_\varepsilon^2} [y_t - (m_{t-1} + b_{t-1})] \\ b_{t-1} + \frac{r_{12(t)}}{r_{11(t)} + \sigma_\varepsilon^2} [y_t - (m_{t-1} + b_{t-1})] \end{bmatrix} \\ C_t &= \begin{bmatrix} c_{11(t)} & c_{12(t)} \\ c_{21(t)} & c_{22(t)} \end{bmatrix} = \begin{bmatrix} \frac{r_{11(t)}\sigma_\varepsilon^2}{r_{11(t)} + \sigma_\varepsilon^2} & \frac{r_{12(t)}\sigma_\varepsilon^2}{r_{11(t)} + \sigma_\varepsilon^2} \\ \frac{r_{12(t)}\sigma_\varepsilon^2}{r_{11(t)} + \sigma_\varepsilon^2} & r_{22(t)} - \frac{r_{12(t)}^2}{r_{11(t)} + \sigma_\varepsilon^2} \end{bmatrix} \end{aligned}$$

لمزيد من المعلومات أنظر [الحمداني، ١٩٩٦]

إن أساسيات التقدير التعاقبي (Recursive Estimation) جرى تفصيلها لأول مرة من قبل العالم (Kalman) في عام ١٩٦٠، حيث وضع حلاً لعملية التقدير التعاقبي للحصول على المقدر الأمثل. وفي الفصل الآتي سنهتم بدراسة هذا الأسلوب فضلاً عن توضيح بعض المشاكل التي تواجهها وطرائق حلها والتغلب عليها.

# الفصل الثاني

مرشح كالمن

ومشكلة التباعد في مصفوفة التباين والتباين المشترك

*Kalman Filter*

*& Divergence Problem in Covariance Matrix*

## الفصل الثاني

### مرشح كالمن ومشكلة التباعد في مصفوفة التباين والتباين المشترك

#### ١-٢- مقدمة:

إن الغاية الأساسية من الترشيح هي الحصول على المقدر الأمثل (Optimal Estimator). وكانت الأساسيات التي بنيت عليها نظرية التقدير عند ظهورها للمرة الأولى تعتمد على جعل دالة تباين الخطأ أقل ما يمكن، وكما هو معلوم فإن طريقة المربعات الصغرى وجدت لأول مرة على يد العالم (Gauss) عام ١٧٩٥. ومنذ ذلك الحين تتابعت الدراسات الموسعة في هذا المجال، وخصوصاً فيما يتعلق بالتقدير الخطي الأمثل (Optimal Linear Estimation) وقد ظهرت أولى الدراسات الرائدة في مسألة الترشيح في مطلع الأربعينات من القرن الماضي، حيث قام كلٌّ من (Kolmogorov) عام ١٩٤١، و (Wiener) عام ١٩٤٢ بشكل مستقل بتطوير تقنية لإيجاد تقدير خطي ذي أقل متوسط مربع خطأ (Linear Minimum Mean-Square Error Estimation) والذي لقي اهتماماً بالغاً وكان له الأثر الكبير لاحقاً في تطوير فكرة مرشح كالمن. [Sorenson, 1970]، ثم توج [Wiener, 1949] دراسته تلك بنتائج أخرى تمخضت عن وضع خوارزمية تعاقبية (Recursive Algorithm) لإيجاد المقدر الخطي الأمثل، وكانت هذه الخوارزمية مقيدة بشروط منها أن المشاهدة المدروسة ذات بُعد أحادي (Scalar) فضلاً عن المعلمة، وأن البيانات شبه غير منتهية ( $t_0 \rightarrow -\infty$ ) فضلاً عن ذلك فإن العملية تنسم بالاستقرار (Stationary Process). وفي عام ١٩٥٠ قام (Zadeh & Ragazzini) بتطوير النتائج التي توصل إليها (Wiener) وذلك بجعلها أكثر عمومية وذلك من حيث جعل البيانات منتهية ( $-\infty < t_0$ ) وشمولها للعمليات التي تنسم بعدم الاستقرار (Non-Stationary Processes). تلا ذلك دراسات أخرى في المجال نفسه نذكر منها على سبيل المثال لا الحصر الدراسة التي قام بها Dolph & Woodbury في عام ١٩٥٢ و Darlington في عام 1959 والتي لم تتعد النتائج التي توصل إليها Zadeh & Ragazzini. [Kailath, 1974]

ومع كثرة الدراسات في هذا المجال فقد بقيت عقيمة نوعاً ما عن تفسيرها لظواهر عديدة ظهرت خصوصاً في أواخر الخمسينات بعد دخول عصر الفضاء وظهور مشكلة تحديد مدار القمر الصناعي (Satellite Orbit) من حيث السرعة (Velocity) كمعلمة أولى والتعجيل (Acceleration) كمعلمة ثانية، وذلك للأسباب الآتية:

١. التعقيد الشديد، ففي كثير من الأحيان تتطلب حلول مساعدة في المعادلات التفاضلية والجبرية، وحساب جذور متعددات الحدود (Polynomials).

٢. يوجد صعوبة وتعقيد في العمليات الحسابية عند أخذ مشاهدة جديدة.

٣. عدم إمكانية التكيف بشكل ملائم مع حالة تعدد المعلمات والمشاهدات.

في تلك الأثناء تمكن (Swerling) في عام ١٩٥٩ من وضع أول خوارزمية تعاقبية (Recursive Algorithm) لحل المشاكل التي تهتم بالنماذج ذات المعلمات المتعددة، فضلاً عن المشاكل الأخرى، ثم قام [Kalman, 1960] بوضع خوارزمية تعاقبية أكثر تعقيداً من تلك التي وضعها (Swerling) لكنها تميزت بكونها أول خوارزمية تهتم بدراسة مشاكل تقدير الحالة الحركي (Dynamical State Estimation Problems)، والتي يسرت فهم وتحليل العديد من المشاكل التطبيقية. وقد كانت الخوارزمية التي وضعها (Kalman) تهتم بدراسة تقدير الحالة الحركي في الأزمنة المتقطعة (Discrete Time) أي ( $t = 1, 2, \dots$ ). ثم قام [Kalman & Bucy, 1961] بوضع خوارزمية لتقدير الحالة الحركي في الأزمنة المستمرة (Continuous Time) وهو ما عرف لاحقاً بـ (Kalman-Bucy Filter). [Kailath, 1974]

بعد ذلك تتابعت الدراسات النظرية والتطبيقية وبشكل موسع جداً على ما سمي لاحقاً بنظرية مرشح كالمن (Kalman Filter Theory)، تضمنت تطوير خوارزميات محدثة لخوارزمية مرشح كالمن الاعتيادي (Ordinary Kalman Algorithm) وذلك للتغلب على المشاكل التي تواجه المستخدم لهذه الخوارزمية، مثل مشكلة النماذج غير الخطية، أو عندما يكون تباين الأخطاء غير معلوم، أو عندما تكون الأخطاء العشوائية مترابطة، انظر [الحمداي، ١٩٩٦]، وكذلك بعض المشاكل العددية التي تحصل في مصفوفة التباين.

ويمكن تعريف مرشح كالمن بأنه خوارزمية تعاقبية لإيجاد مقدر أمثل لمتغير الحالة أو المعلمة، أي يكون تباين خطأ التقدير أقل ما يمكن. وتكمن المشكلة التي تطرق إليها كالمن في إيجاد المقدر الأمثل للمعلمة ( $\theta_r$ ) وقد أطلق كالمن مسميات مختلفة على التقدير وذلك تبعاً لقيمة ( $\tau$ ) كمؤشر للزمن المطلوب التقدير عنده مقارنة مع الزمن الحاضر ( $t$ )، فإذا كانت ( $\tau < t$ ) فتعرف المشكلة بالتمهيد (Smoothing)، وإذا كانت ( $\tau = t$ ) فيطلق على المشكلة مسمى الترشيح (Filtering)، وأخيراً إذا كانت ( $\tau > t$ ) فتعرف المشكلة عندئذٍ بالتنبؤ (Prediction)، والشكل (٢-١) يوضح الفرق بين الحالات الثلاث.

وقد اعتمد كالمن في إيجاده للمقدر الخطي الأمثل على فكرة الإسقاط المتعامد (Orthogonal Projections) أو ما يعرف بمبدأ التعامد (Orthogonal principle)، أي باستخدام طريقة المربعات الصغرى، والذي يمكن تفصيله بشكل مبسط كالاتي:

لو فرضنا أنه لدينا النموذج الخطي الحركي الآتي:

$$y_t = \theta_t + \varepsilon_t$$

$$\theta_t = \theta_{t-1} + \delta_t$$

وليكن  $(\hat{\theta}_t)$  مقدر لمتغير الحالة أو المعلمة  $(\theta_t)$  في الزمن  $(t)$ ، فعندئذٍ إذا كان:

$$E[(\hat{\theta}_t - \theta_t)y_t] = 0$$

حيث أن  $(\hat{\theta}_t - \theta_t)$  هو خطأ التنبؤ.

فإن  $(\hat{\theta}_t)$  يكون أفضل مقدر خطي لـ  $(\theta_t)$ . وقد توصل كالمن إلى أن المقدر الخطي الأمثل هو التوقع الشرطي لـ  $(\theta_t)$  مشروطاً بكافة المعلومات المتوافرة حتى الزمن  $(t)$  أي:

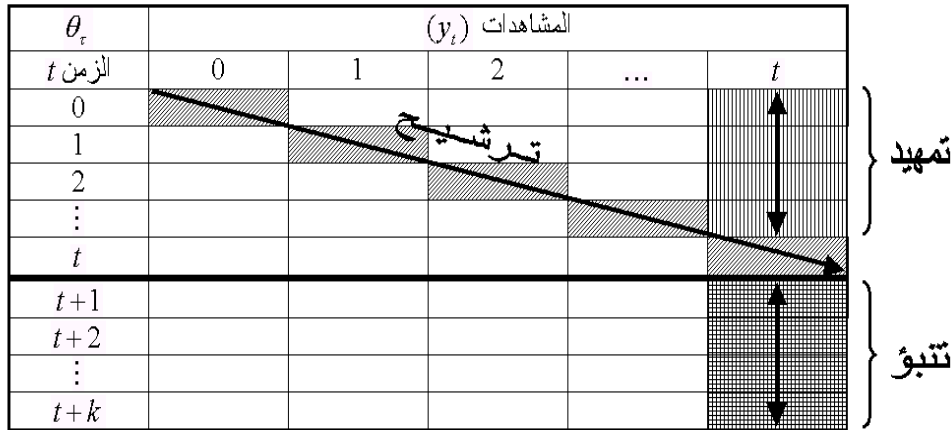
$$E(\theta_t | D_t)$$

وأن هذا المقدر يمكن الحصول عليه إذا توافر أحد الشرطين الآتيين:

١. إذا توزع كل من متغير الحالة  $(\theta_t)$  والمشاهدات  $(y_t)$  وفق التوزيع الطبيعي.

٢. إذا كانت دالة الخسارة هي دالة الخسارة التربيعية (Quadratic Loss Function)، أي

$$\text{أن: } l(\hat{\theta}_t, \theta_t) = (\hat{\theta}_t - \theta_t)^2 \text{ [Kalman, 1960]}$$



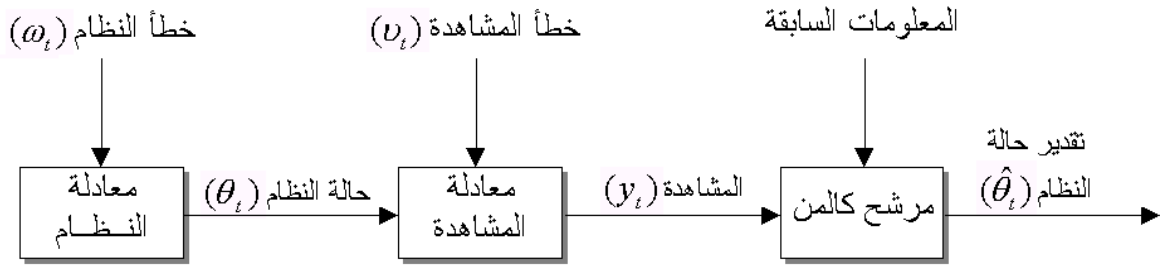
الشكل (١-٢)

الفرق بين التمهيد والترشيح والتنبؤ. [Snyder & Forbes, 2002]

وإذا ما قارنا المقدر الخطي الأمثل الذي توصل إليه كالمن بطريقة المربعات الصغرى التعاقبية مع مقدر بيز ذي أقل متوسط مربع خطأ (MMSE) عند إيجاده بطريقة التحليل المتسلسل (Sequential Analysis) نلاحظ أن مقدر (MMSE) هو نفسه المقدر الذي توصل إليه كالمن، وهذا بالضبط ما توصل إليه الباحثان [Harrison & Stevens, 1976]. إن استخدام مبرهنة بيز في

عملية تكوين مرشح كالمن تعطي سلاسة وسهولة عاليتين في العمليات الحسابية التي يتطلبها إيجاد مرشح كالمن مقارنة مع الأسلوب الذي وضعه كالمن للمرة الأولى عام ١٩٦٠ والذي يتسم بالتعقيد بعض الشيء. ومن ابرز من استخدم أسلوب بيز في تكوين مرشح كالمن [Maybeck, 1979] و [Meinhold & Singpurwalla, 1983] و [West & Harrison, 1989] و [Calder, et al. 2003] وغيرهم آخرون.

والشكل أدناه يوضح باختصار عملية التقدير الحركي التي وضعها كالمن لإيجاد المقدر الأمثل.



الشكل (٢-٢)

خطوات الخوارزمية التفاعلية التي وضعها كالمن لإيجاد المقدر الأمثل. [Gelb, 1974].

## ٢-٢- بعض التطبيقات التي يستخدم فيها مرشح كالمن:

لقد استُخدم مرشح كالمن على نطاق واسع في العديد من الحقول التطبيقية، وخصوصاً تلك التي تطورت ونشأت بعد الثورة العلمية التي حدثت في مجال هندسة الاتصالات والكمبيوتر وعلم الفضاء، ومن هذه الحقول التطبيقية:

١. التطبيقات الهندسية: مثل معالجة الإشارة الصادرة عن الأقمار الصناعية أو الصادرة عن التعقيب بواسطة الرادار (Radar Tracking)، وفي أنظمة الملاحة (Navigation Systems) الخاصة بالمركبات الفضائية والطائرات والسيارات.
٢. التطبيقات الحاسوبية: مثل معالجة الصورة، وتحديد الرسوم الفورية Real-Time Graphics، وفي المعالجة الصوتية.
٣. التطبيقات الاقتصادية والإحصائية: مثل التنبؤ بالمؤشرات الاقتصادية، وفي مجال السيطرة الإحصائية ونظرية السيطرة (Control Theory).

أنظر [Welch, & Bishop, 2001].

## ٣-٢- مرشح كالمن ومبرهنة بيز:

كما سبق وأن ذكرنا، فإن استخدام مبرهنة بيز لإيجاد المقدّر الخطي الأمثل الذي وضعه كالمن تعطي سلاسة وسهولة في العمليات الحسابية، وأن الباحثين (Harrison & Stevens) هما أول من توصل إلى تكوين مرشح كالمن للنماذج الخطية الحركية باستخدام التحليل المتسلسل مع إحصاء بيز (Bayesian Sequential Analysis) عام ١٩٧٦، لذا سنستعرض بإيجاز جزءاً من النتائج التي توصلنا إليها في هذا المجال.

فلو فرضنا أنه لدينا النموذج الخطي الحركي المعرفّ بمنظومة المعادلات  $(1, \epsilon)$ ، فإن تكوين مرشح كالمن لهذا النموذج عند الزمن  $(t)$  يتلخص في الخطوات الآتية:

١. التوزيع اللاحق (Posterior Distribution) لمتجه الحالة عند الزمن  $(t-1)$  يكون

توزيعاً طبيعياً متعدداً بـ  $n$  من المتغيرات له متجه متوسط  $M_{t-1}$  ومصفوفة تباين  $C_{t-1}$ ، أي أن:

$$(\theta_{t-1} | D_{t-1}) \sim N_n(M_{t-1}, C_{t-1})$$

٢. التوزيع الأولي (Prior Distribution) لمتجه الحالة عند الزمن  $(t)$  يكون توزيعاً طبيعياً

بمتجه متوسط  $(G_t M_{t-1})$  ومصفوفة تباين  $(R_t)$ ، حيث:

$$R_t = G_t C_{t-1} G_t' + W_t \quad \dots\dots\dots(2.1a)$$

أي أن:

$$(\theta_t | D_{t-1}) \sim N_n(G_t M_{t-1}, R_t)$$

٣. التوزيع الشرطي لمتجه المشاهدات  $(Y_t | \theta_t)$  والذي يمثل دالة الإمكان عند الزمن  $(t)$  يكون

توزيعاً طبيعياً بمتجه متوسط  $(F_t \theta_t)$  ومصفوفة تباين  $(V_t)$ ، أي أن:

$$(Y_t | \theta_t) \sim N_m(F_t \theta_t, V_t)$$

٤. التوزيع اللاحق لمتجه الحالة عند الزمن  $(t)$  هو توزيع طبيعي بمتجه متوسط  $(M_t)$  ومصفوفة

تباين  $(C_t)$ ، حيث أن:

$$M_t = G_t M_{t-1} + K_t (Y_t - F_t G_t M_{t-1}) \quad \dots\dots\dots(2.1b)$$

$$C_t = R_t - K_t F_t R_t \quad \dots\dots\dots(2.1c)$$

$$K_t = R_t F_t' (F_t R_t F_t' + V_t)^{-1} \quad \dots\dots\dots(2.1d)$$

حيث أن  $(K_t)$  تمثل مصفوفة ربحية المرشح (Filter Gain Matrix) وهي ذات بُعد  $(n \times m)$ .

إن منظومة المعادلات (٢,١) أعلاه تعرف بخوارزمية مرشح كالمن  
 Kalman Filter Algorithm [Melsa & Cohn, 1978]، وتعرف أيضاً بخوارزمية الترشيح  
 الأمامي Forward Filtering Algorithm [Calder, et al. 2003].  
 إن المقدر الخطي الأمثل نحصل عليه بعد عدة مراحل زمنية، ويعتمد عدد تلك المراحل الزمنية  
 على وصولنا لحالة الاستقرار (Steady State) والمؤشر العلمي الدقيق هو استقرار مصفوفة ربحية  
 المرشح أي:

$$K_{\tau} = K_{\tau-1} \quad , \tau = t^* + 1, t^* + 2, \dots$$

حيث ( $t^*$ ) هو الزمن الذي يحصل فيه الاستقرار لقيم ( $K_t$ ).

إن الحالات التي تكون فيها المصفوفة  $V_t$  ثابتة لجميع الفترات الزمنية فإن حالة الاستقرار في  
 ربحية المرشح  $K_t$  تحصل عندما تتقارب مصفوفة التباين  $C_t$ ، أي عندما:

$$C_T = C_{T-1} \quad , T = t^* + 1, \dots$$

وبالتالي فإن:

$$R_T = R_{T-1} \quad , T = t^* + 1, \dots$$

وهذا بالتأكيد يؤدي إلى استقرار ربحية المرشح (لاحظ المعادلة (2.1d)).

**مبرهنة (٢-١):**

إن مصفوفة التباين  $C_t$  يمكن صياغتها بالصيغة الآتية:

$$C_t = (I_n - K_t F_t) R_t (I_n - K_t F_t)' + K_t V_t K_t'$$

**البرهان:**

من المعادلة (2.1c) فإن مصفوفة التباين تساوي:

$$C_t = R_t - K_t F_t R_t$$

نضيف ونطرح الحد ( $K_t F_t R_t$ ) من طرف اليمين:

$$C_t = R_t - 2K_t F_t R_t + K_t F_t R_t$$

نضرب الحد الثالث من الطرف الأيمن بالحد ( $F_t R_t F_t' + V_t$ ) ومعكوسه:

$$C_t = R_t - 2K_t F_t R_t + K_t (F_t R_t F_t' + V_t) (F_t R_t F_t' + V_t)^{-1} F_t R_t$$

ومن المعادلة (2.1d) فإن:

$$\begin{aligned} C_t &= R_t - 2K_t F_t R_t + K_t (F_t R_t F_t' + V_t) K_t' \\ &= R_t - 2K_t F_t R_t + K_t F_t R_t F_t' K_t' + K_t V_t K_t' \\ &= R_t (I_n - 2K_t F_t + K_t F_t F_t' K_t') + K_t V_t K_t' \\ \therefore C_t &= (I_n - K_t F_t) R_t (I_n - K_t F_t)' + K_t V_t K_t' \quad \blacksquare \end{aligned}$$



في خوارزمية مرشح كالمن التعااقبية، يلاحظ أن متوسط التوزيع اللاحق ( $M_t$ ) في (2.1b) هو عبارة عن تركيب خطي (Linear Combination) مكون من حدين، الحد الأول هو معدل التوزيع الأولي ( $G_t M_{t-1}$ ) في الزمن ( $t$ )، بينما يمثل الحد الثاني حد تصحيحي (Correction Term) وهو الفرق الموزون بين المشاهدات الحقيقية ( $Y_t$ ) والقيم التنبؤية ( $\hat{Y}_t = F_t G_t M_{t-1}$ ) يعرف هذا الفرق بالخطأ التنبؤي (Prediction Error) ويرمز له بـ ( $e_t$ ). [Harrison & Stevens, 1976]، وأن الوزن هو عبارة عن المصفوفة ( $K_t$ ) أي مصفوفة ربحية المرشح، حيث تقوم هذه المصفوفة بدور عامل التصحيح (Correction Factor). وبذلك يمكن كتابة المعادلة (2.1b) كالآتي:

$$M_t = G_t M_{t-1} + K_t e_t$$

**مبرهنة (٢-٢):**

إن معيار اختيار المصفوفة ( $K_t$ ) هو تقليلها للمجموع الموزون لعناصر القطر الرئيسي لمصفوفة التباين ( $C_t$ ).

**البرهان:**

من المبرهنة (١-٢) فإن مصفوفة التباين ( $C_t$ ) يمكن صياغتها بالصيغة الآتية:

$$C_t = (I_n - K_t F_t) R_t (I_n - K_t F_t)' + K_t V_t K_t'$$

ولغرض إثبات أن المصفوفة ( $K_t$ ) تجعل من التباين أقل ما يمكن نشق المصفوفة ( $C_t$ ) بالنسبة لـ ( $K_t$ ) ونساوي بالصفري، وكالآتي:

$$\frac{\partial C_t}{\partial K_t} = -2(I_n - K_t F_t) R_t F_t' + 2K_t V_t = 0$$

$$\Rightarrow K_t F_t R_t F_t' + K_t V_t - R_t F_t' = 0$$

$$\Rightarrow K_t (F_t R_t F_t' + V_t) = R_t F_t'$$

$$\therefore K_t = R_t F_t' (F_t R_t F_t' + V_t)^{-1} \quad \blacksquare$$

وبالنظر إلى المعادلة (2.1d) نلاحظ أن المصفوفة ( $K_t$ ) تعتمد على مصفوفة تباين خطأ المشاهدة ( $V_t$ ) ومصفوفة تباين التوزيع الأولي للحالة ( $R_t$ )، فإذا كانت مصفوفة تباين خطأ المشاهدة ( $V_t$ ) قريبة من الصفر فإن مصفوفة ربحية المرشح ( $K_t$ ) ستعطي وزناً كبيراً لخطأ التنبؤ، وتحديدًا تكون:

$$\lim_{V_t \rightarrow 0} K_t = F_t^{-1}$$

ومن جهة أخرى إذا اقتربت مصفوفة تباين التوزيع الأولي للحالة ( $R_t$ ) من الصفر فإن مصفوفة ربحية المرشح ( $K_t$ ) ستعطي وزناً قليلاً لخطأ التنبؤ، وتحديداً تكون:

$$\lim_{R_t \rightarrow 0} K_t = 0$$

وهذا يُفسر بأنه إذا كانت المصفوفة ( $V_t$ ) تقترب من الصفر فإنه بالتأكيد سوف تزداد ثقتنا بالمشاهدة الحقيقية ( $Y_t$ ) أكثر فأكثر، مقابل معدل التنبؤ ( $F_t G_t M_{t-1}$ ). بينما إذا اقتربت مصفوفة تباين التوزيع الأولي للحالة ( $R_t$ ) من الصفر يعني أن الثقة بالمشاهدة الحقيقية ( $Y_t$ ) تبدأ بالنقصان، بينما تزداد ثقتنا بمعدل التنبؤ ( $F_t G_t M_{t-1}$ ) أكثر فأكثر. [Welch, & Bishop, 2001]

كما يلاحظ أنه إذا كانت القيم الذاتية (Eigenvalues) للمصفوفة ( $W_t$ ) أكبر مقارنة بالقيم الذاتية للمصفوفة ( $V_t$ ) فإننا سوف نصل إلى حالة الاستقرار والمقدر الأمثل بسرعة، وذلك لأن الجزء غير المحدد (Undeterministic) في معادلة النظام يكون أكبر مقارنة بالدقة المتوافرة في معادلة المشاهدة، وذلك يؤدي لأن يعتمد التقدير الجديد لـ ( $\theta_t$ ) على قيمة المشاهدة الجديدة أكثر من اعتماده على التقدير السابق. [Maybeck, 1979]

ومن الجدير بالذكر أن العلاقات التعاقبية (2.1a) و (2.1c) و (2.1d) لا تعتمد في حسابها بشكل مباشر على قيم المشاهدات ( $Y_t$ )، ومع أن مصفوفة التباين ترتبط إحصائياً بالمصفوفتين ( $F_t$ ) و ( $V_t$ ) الخاصتين بالملاحظات، غير أن هاتين المصفوفتين ليستا دالتين للملاحظات المحددة. ولذلك يمكن حساب المصفوفات الثلاث ( $R_t$ ) و ( $K_t$ ) و ( $C_t$ ) مسبقاً قبل ظهور المشاهدة ( $y_t$ ). وبما أن ( $R_t$ ) و ( $C_t$ ) هما مصفوفتي تباين التوزيع الأولي والتوزيع اللاحق للحالة ( $\theta_t$ ) على التوالي، فإن الحساب المسبق لهاتين المصفوفتين يمكننا من أخذ فكرة عن مدى الكلفة التي سنتحملها باستبدالنا المقدر السابق بالمقدر اللاحق ( $M_t$ ) مقارنة مع دقة القياسات التي نستدل عليها من المصفوفة ( $V_t$ ) لضمان الحصول على مقدر للحالة بالمواصفات المطلوبة بأقل كلفة ممكنة. [Maybeck, 1979]

من جهة أخرى فإن الحساب المسبق للمصفوفة ( $K_t$ ) يمكننا من الاستدلال عن مدى أهمية عملية الترشيح، حيث عندما تكون  $K_t \cong [0]$  فإن ذلك يعني أن قيمة المشاهدة الجديدة لا تؤثر في عملية الترشيح، والعكس صحيح. [الحمداني، ١٩٩٦].

### مبرهنة (٣-٢):

إن المصفوفة ( $K_t$ ) يمكن صياغتها بصيغة أخرى توفر سهولة في العمليات الحسابية، وهي:

$$K_t = C_t F_t V_t^{-1}$$

البرهان:

من المعادلة (2.1d) فإن:

$$\begin{aligned} K_t &= R_t F_t' (F_t R_t F_t' + V_t)^{-1} \\ &= C_t C_t^{-1} R_t F_t' (F_t R_t F_t' + V_t)^{-1} \end{aligned}$$

وبالتعويض عن  $C_t^{-1}$  بما يساويها في المعادلة (1,10):

$$\begin{aligned} K_t &= C_t (R_t^{-1} + F_t V_t^{-1} F_t) R_t F_t' (F_t R_t F_t' + V_t)^{-1} \\ &= (C_t F_t' + C_t F_t V_t^{-1} F_t R_t F_t') (F_t R_t F_t' + V_t)^{-1} \\ &= C_t F_t V_t^{-1} (V_t + F_t R_t F_t') (F_t R_t F_t' + V_t)^{-1} \\ K_t &= C_t F_t V_t^{-1} \quad \blacksquare \end{aligned}$$

## ٢-٤ - أنواع مرشح كالمن (Kalman Filter Types):

إن الفروض الواجب توافرها لإيجاد المقدر الخطي الأمثل باستخدام خوارزمية مرشح كالمن الاعتيادي (Ordinary Kalman Filter) قد لا تتوافر دائماً، بل إنها غالباً ما تكون غير متوافرة تطبيقياً، لذلك قام العلماء والباحثون بوضع عدة خوارزميات مشابهة لخوارزمية مرشح كالمن الاعتيادي مع بعض التغييرات والإضافات وذلك تبعاً للمشكلة المراد معالجتها.

وبدئ ذي بدئ فإن مرشح كالمن يقسم بشكل عام - بما فيه النوع الاعتيادي- إلى نوعين رئيسيين:

**الأول:** مرشح كالمن للتباين والتباين المشترك ويعرف بـ (Covariance Kalman Filter)، ويعتمد على مصفوفة التباين والتباين المشترك.

**الثاني:** مرشح كالمن للمعلومات ويعرف بـ (Information Kalman Filter)، ويعتمد على

مصفوفة المعلومات، ويعرف أيضاً بمرشح كالمن المعكوس (Inverse Kalman Filter).

ومن ثم يمكن تقسيم مرشح كالمن تحت هذين النوعين إلى أنواع مختلفة وذلك حسب المشكلة التي

تواجهنا، ومن هذه الأنواع:

١. مرشح كالمن التكيفي (Adaptive Kalman Filter): ويستخدم عندما تكون تباينات الأخطاء

غير معلومة، حيث يجري تقديرها. [الحمداني، ١٩٩٦].

٢. مرشح كالمن الجذر التربيعي (Square Root Kalman Filter): ويستخدم هذا النوع عندما

تحدث مشاكل عددية في مصفوفة التباين والتباين المشترك بسبب أخطاء التدوير

(Round-off Errors). [Bellantoni, & Dodge, 1967]

٣. مرشح كالمن الموسع (Extended Kalman Filter): ويستخدم عندما تكون النماذج الحركية المدروسة غير خطية، حيث يتم تحويلها إلى نماذج خطية مقربة ثم يتم إيجاد المقدر لها. [الرسام، ٢٠٠٤]

٤. مرشح كالمن عديم الأثر (Unscented Kalman Filter): ويستخدم هذا النوع أيضاً عندما تكون النماذج المدروسة غير خطية ويفضل في بعض الأحيان على مرشح كالمن الموسع. [Wan, et al. 2000].

وسيكون اهتمامنا في هذه الدراسة بخوارزمية مرشح كالمن الجذر التربيعي، وقبل البدء بتوضيح المشكلة وطريقة التغلب عليها نورد بعض الفقرات المهمة التي سنحتاجها في اشتقاقنا لهذه الخوارزمية.

## ٥-٢- الصيغة القطرية للمصفوفة (The Diagonal Form of a Matrix):

افرض أن المصفوفة ( $A$ ) مربعة وذات بعد ( $n \times n$ )، وأن لهذه المصفوفة ( $n$ ) من المتجهات الذاتية المستقلة خطياً (Linearly Independent Eigenvectors) يرمز لها بـ  $(e_i : i = 1, 2, \dots, n)$ ، فعندئذٍ إذا اختيرت هذه المتجهات لتكون أعمدة المصفوفة ( $E$ )، فإن  $(E^{-1}AE)$  عبارة عن مصفوفة قطرية ( $\Lambda$ ) تمثل عناصر القطر فيها القيم الذاتية (Eigenvalues) للمصفوفة ( $A$ )، أو ما يعرف بالجذور المميزة للمصفوفة (Characteristic Roots) ويرمز لها بـ  $(\lambda_i : i = 1, 2, \dots, n)$ ، أي أن:

$$E^{-1}AE = \Lambda$$

$$\begin{bmatrix} \vdots & \vdots & \dots & \vdots \\ e_1 & e_2 & \dots & e_n \\ \vdots & \vdots & \dots & \vdots \end{bmatrix}^{-1} \times \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \times \begin{bmatrix} \vdots & \vdots & \dots & \vdots \\ e_1 & e_2 & \dots & e_n \\ \vdots & \vdots & \dots & \vdots \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

أنظر: [Strang, 1980]

## ٦-٢- المبرهنة الطيفية (The Spectral Theorem):

لأي مصفوفة حقيقية مربعة ومتماثلة تكون القيم الذاتية لتلك المصفوفة حقيقية أيضاً، وتكون المتجهات المميزة متعامدة - إذا كانت المصفوفة غير متماثلة فإن المتجهات المميزة لا تكون متعامدة-، فإذا تم تحويل هذه المتجهات إلى الصيغة الطبيعية (Normalize) ووضع المتجهات الناتجة في المصفوفة ( $E$ )، فإن المصفوفة ( $E$ ) تكون متعامدة وذات صيغة طبيعية متعامدة (Orthonormal). أي أن:

$$E^{-1} = E'$$

وبذلك تكون الصيغة القطرية للمصفوفة (A) كالآتي:

$$E^{-1}AE = E'AE = \Lambda$$

وبضرب الطرفين ضرباً مسبقاً بالمصفوفة (E) وضرباً لاحقاً بالمصفوفة (E')، يكون:

$$A = E\Lambda E'$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ e_1 & e_2 & \cdots & e_n \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix} \times \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \times \begin{bmatrix} \cdots & e_1 & \cdots \\ \cdots & e_2 & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & e_n & \cdots \end{bmatrix}$$

حيث أن:

$$EE' = E'E = EE^{-1} = E^{-1}E = I$$

إن الصيغة الأخيرة تعرف بالتحليل الطيفي للمصفوفة Spectral Decomposition of a Matrix  
أنظر: [Strang, 1980]

## ٧-٢- الجذر التربيعي للمصفوفة (The Square Root of a Matrix)

إذا كانت المصفوفة (A) مربعة، فعندئذ يكون الجذر التربيعي للمصفوفة (A) - يرمز له بـ  $(A^{1/2})$  -

هو أي مصفوفة تحقق المعادلة الآتية:

$$A = (A^{1/2})(A^{1/2})' \quad \dots\dots\dots(2.2)$$

وتوجد أكثر من طريقة لإيجاد الجذر التربيعي للمصفوفة، وسنوضح هنا طريقتين لإيجاد المصفوفة التي تحقق المعادلة (٢،٢).

### ١-٧-٢- الجذر التربيعي للمصفوفة باستخدام التحليل الطيفي:

كما ذكرنا سابقاً فإن التحليل الطيفي للمصفوفة (A) يكون كالآتي:

$$A = E\Lambda E'$$

وبما أن المصفوفة (Λ) قطرية، فإن الجذر التربيعي لها يمكن إيجاده بأخذ الجذر التربيعي

لعناصرها، أي أن:

$$\Lambda^{1/2} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n} \end{bmatrix}$$

إذن:

$$E\Lambda E' = E\Lambda^{1/2}\Lambda^{1/2}E'$$

وبما أن المصفوفة ( $E$ ) متعامدة، فإن:

$$A = (E\Lambda^{1/2}E')(E\Lambda^{1/2}E')'$$

وبالمقارنة مع المعادلة (٢,٢):

$$A = (A^{1/2})(A^{1/2})'$$

حيث أن:

$$A^{1/2} = E\Lambda^{1/2}E'$$

إن الجذر التربيعي للمصفوفة عند إيجاده بهذه الطريقة لا يكون فريداً (Not Unique)، وذلك لأنه إذا كانت المصفوفة ( $A^{1/2}$ ) تمثل جذراً تربيعياً للمصفوفة ( $A$ )، وكانت ( $T$ ) مصفوفة متعامدة ( $TT' = I$ )، فعندئذٍ تكون المصفوفة ( $A^{1/2}T$ ) جذراً تربيعياً أيضاً للمصفوفة ( $A$ ).

[Bierman, 1977]

## ٢-٧-٢ - تحليل چولسكي (Cholesky Decomposition):

إذا كانت المصفوفة ( $A$ ) متماثلة وأكيدة الإيجابية، فعندئذٍ يمكن تحليلها إلى مصفوفتين مثلثيتين، الأولى مصفوفة مثلثية سفلى (Lower Triangular Matrix) والثانية مصفوفة مثلثية عليا (Upper Triangular Matrix) وتمثل المصفوفة الثانية مبدول المصفوفة الأولى، أي أن:

$$A = LL'$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix} \times \begin{bmatrix} l_{11} & l_{21} & \cdots & l_{n1} \\ 0 & l_{22} & \cdots & l_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & l_{nn} \end{bmatrix}$$

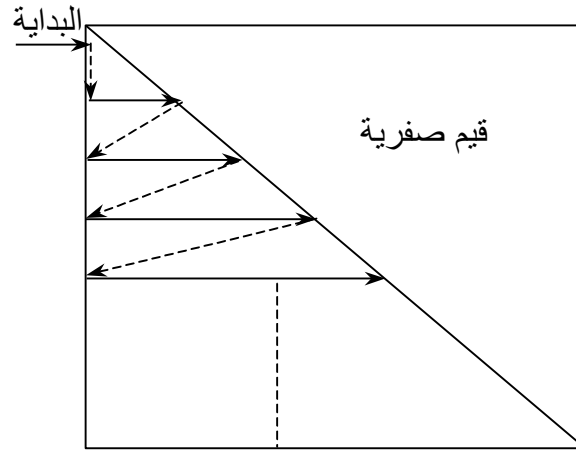
حيث أن  $L$  تمثل المصفوفة المثلثية السفلى.

ويمكن إيجاد عناصر المصفوفة المثلثية السفلى بشكل متعاقب صف بعد صف كالآتي:  
إذا كانت  $(i = 1, 2, \dots, n)$  فإن:

$$l_{ij} = \begin{cases} \left( \frac{1}{l_{jj}} \right) \times \left[ a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk} \right] & , j = 1, 2, \dots, i-1 \\ \left( a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2 \right)^{1/2} & , j = i \\ 0 & , j > i \end{cases} \dots\dots\dots (2.3)$$

والشكل (١-٢) يوضح طريقة مسح (Scanning) المصفوفة (A) لإيجاد المصفوفة المثلثية السفلى. وبمقارنة تحليل چولسكي مع المعادلة (٢,٢) نستنتج أن المصفوفة المثلثية السفلى تكون جذراً للمصفوفة (A)، أي أن  $(L = A^{1/2})$ .

أنظر [Maybeck, 1979]



الشكل (١-٢)

المراحل المتعاقبة لإيجاد المصفوفة المثلثية السفلى لچولسكي. [Maybeck, 1979]

### توضيح: <

إذا كان لدينا المصفوفة الآتية:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 8 & 2 \\ 3 & 2 & 14 \end{bmatrix}$$

وإذا أردنا إيجاد الجذر التربيعي لهذه المصفوفة باستخدام تحليل چولسكي، فنقوم بإيجاد كل عنصر من عناصر مصفوفة الجذر التربيعي باستخدام (٢,٣) وكالآتي:

$$\begin{aligned} l_{11} &= \sqrt{1} = 1 & l_{12} &= 0 & l_{13} &= 0 \\ l_{21} &= (1/1)[2-0] = 2 & l_{22} &= \sqrt{8-4} = 2 & l_{23} &= 0 \\ l_{31} &= (1/1)[3-0] = 3 & l_{32} &= (1/2)[2-6] = -2 & l_{33} &= \sqrt{14-9-4} = 1 \end{aligned}$$

إذن:

$$L = A^{1/2} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 3 & -2 & 1 \end{bmatrix}$$

وبسهولة يمكن استنتاج أن:

$$LL' = (A^{1/2})(A^{1/2})' = A$$

## ٨-٢- مرشح كالمن الجذر التربيعي:

### ١-٨-٢- التباعد (Divergence):

بسبب أخطاء النمذجة أو الظروف العددية غير الملائمة أو أخطاء التدوير... الخ، فإن خوارزمية مرشح كالمن يمكن في بعض الأحيان أن تعطي نتائج (مقدرات) أقل دقة بكثير مما نتوقعه من خلال الاستدلال بمصفوفة التباين من حيث كونها صغيرة بشكل غير طبيعي، فتكون الخوارزمية غير مستقرة (كأن يكون مقدر الحالة يزداد خلال الزمن بشكل أسي متجاوزاً جميع القيود)، كما أن احتمال الإخفاق في الحصول على نتائج وارد بسبب فقدان خاصية أكيدة الإيجابية Positive Definite لمصفوفة التباين  $C_T$ ، إن هذه الحالات تعرف بشكل عام بالتباعد.

[Anderson, & Moore, 1979]

ويمكن أن نعزو مشكلة التباعد في بعض الحالات لأحد الأسباب الآتية:

١. التباعد بسبب استخدام معلومات ابتدائية غير صحيحة أو بسبب النمذجة غير الدقيقة.
٢. التباعد بسبب وجود نماذج لا خطية.
٣. التباعد بسبب أخطاء التدوير.

أنظر [Bierman, 1977]

وتؤدي مشكلة التباعد هذه عادةً إلى إظهار قيم صغيرة بشكل غير طبيعي لمصفوفة التباين ( $C_T$ ) ومصفوفة ربحية المرشح ( $K_T$ )، وبالمقابل مقدار كبير نسبياً للمقدر ( $M_T$ ). علاوةً على ذلك، فإن أخطاء التنبؤ التعاقبية ( $e_T's$ ) لا تكون مستقلة في هذه الحالة، أي تكون مترابطة، ولذلك فإن الفروض التي بنيت عليها خوارزمية مرشح كالمن الاعتيادي لا تتوافر.

وقد أظهرت الخبرة [Andrews, 1968] أن الصيغة (2.1c) تقودنا في بعض الأحيان للحصول على مصفوفة تباين لها قيم ذاتية صفرية أو سالبة - وهو ما يعد مستحيلًا نظرياً وعملياً- أي تكون مصفوفة التباين ( $C_T$ ) غير أكيدة الإيجابية - شبه أكيدة الإيجابية أو أكيدة السلبية- مما يؤدي إلى فشل الإجراء أو يؤدي ذلك للحصول على نتائج غير منطقية ومتقلبة (Unstable Results).



وقد وضع الباحثون والخبراء بعض المعالجات للتغلب على هذه المشكلة ومن هذه المعالجات:

١. استخدام تباينات أكبر لخطأ القياسات وخطأ النظام.

٢. استخدام الصيغة الآتية لإيجاد مصفوفة التباين  $(C_t)$ :

$$C_t = (I_n - K_t F_t) R_t (I_n - K_t F_t)' + K_t V_t K_t'$$

واقترحت هذه المعالجة من قبل (Joseph, P.D.) في عام ١٩٦٤ وتعمل هذه الصيغة على

جعل مرشح كالمن مستقرّاً (Stabilized Kalman Filter). [Carlson, 1973].

٣. تقييد مصفوفة التباين  $(C_t)$ .

٤. استخدام خوارزمية أخرى أكثر استقراراً.

أنظر [Bierman, 1977; Kybic, 1998]

وتعد خوارزمية مرشح كالمن الجذر التربيعي أحد الخوارزميات التي تعالج هذه المشكلة، إن الفكرة الأساسية لمرشح كالمن الجذر التربيعي تكمن في تحليل مصفوفة التباين  $(C_t)$  إلى مصفوفتين، تمثل الأولى مصفوفة الجذر التربيعي لمصفوفة التباين وتمثل الثانية مبدول مصفوفة الجذر التربيعي وكالاتي:

$$C_t = (C_t^{1/2})(C_t^{1/2})'$$

حيث أن  $(C_t^{1/2})$  هي مصفوفة الجذر التربيعي للمصفوفة  $(C_t)$ .

إن استخدام مصفوفة الجذر التربيعي لمصفوفة التباين يضمن لنا بقاء مصفوفة التباين لها قيم ذاتية

موجبة (مصفوفة أكيدة الإيجابية). [Morf, & Kailath, 1975]

كما أن استخدام خوارزمية مرشح كالمن الجذر التربيعي يوافر دقة مضاعفة في العمليات الحسابية ويجعل الخوارزمية أكثر ثباتاً بشكل استثنائي. أما إذا استمر فشل الإجراء فهذا يعني أن مرشح كالمن

على الأغلب غير ملائم للحالة المدروسة. [Kybic, 1998]

ويلاحظ أن العدد الشرطي (Condition Number) الخاص بمصفوفة الجذر التربيعي لمصفوفة

التباين يساوي الجذر التربيعي للعدد الشرطي لمصفوفة التباين نفسها، أي:

$$k(C_t^{1/2}) = [k(C_t)]^{1/2}$$

لذا تكون مصفوفة الجذر التربيعي أقل تأثراً بالمشاكل العددية (Numerical Problems).

(أنظر الملحق B)

لقد ابتكر الباحثون العديد من خوارزميات مرشح كالمن الجذر التربيعي في السنوات القليلة التي

تلت ظهور خوارزمية مرشح كالمن الاعتيادي، وتختلف هذه الخوارزميات فيما بينها بطريقة تحليل

مصفوفة التباين وبالصيغة التي يتم من خلالها حساب مصفوفة التباين  $(C_t)$ .

إن إيجاد مصفوفة الجذر التربيعي لمصفوفة التباين باستخدام التحليل الطيفي حيث مصفوفة الجذر تكون غير فريدة (Not Unique) لا يشكل صعوبة أو مشكلة لنا، وتوجد بعض الخوارزميات التي تعتمد على هذا الأسلوب فقط، أما إذا كانت الخوارزمية تتطلب أن تكون مصفوفة الجذر التربيعي فريدة فسوف نستخدم مصفوفة الجذر التربيعي الناتجة من تحليل چولسكي. [Carlson, 1973]

وفي الفقرات الآتية سنوضح بعض أنواع مرشح كالمن الجذر التربيعي.

## ٢-٨-٢- مرشح الجذر التربيعي لبوتر (Potter Square Root Filter):

يعد بوتر (Potter) هو أول من وضع خوارزمية مرشح كالمن الجذر التربيعي في عام (١٩٦٣)، [Carlson, 1973]، لذلك تعرف هذه الخوارزمية بمرشح الجذر التربيعي لبوتر (Potter Square Root Filter)، وكانت هذه الخوارزمية في حينها تعتمد على النموذج الخطي الحركي ذو المشاهدات المفردة ( $m=1$ )، فضلاً عن عدم وجود أخطاء عشوائية في معادلة النظام ( $W_t = 0$ ) أي اعتبار معادلة النظام تامة وهو ما يعرف بـ (Exact System Equation) مما يجعلها مقيدة بهذين الشرطين، ولقد بين بوتر أنه بالإمكان إعادة صياغة خوارزمية مرشح كالمن باستخدام الجذر التربيعي للمصفوفة، ولقد استخدمت هذه الخوارزمية بنجاح في برنامج أبولو الفضائي إبان الستينات. [Bierman, 1977]، وفيما يأتي توضيح لهذه الخوارزمية:  
إذا كان لدينا النموذج الخطي الحركي الآتي:

$$y_t = F_t \theta_t + \varepsilon_t$$

$$\theta_t = G_t \theta_{t-1}$$

حيث أن مكونات هذا النموذج تعرف كما في السابق، مع ملاحظة أن ( $m=1$ )، ( $n > 1$ )

$$. (W_t = 0) ، (V_t = \sigma_\varepsilon^2) ، (\omega_t = 0) ، (v_t = \varepsilon_t)$$

ومن المعادلة (2.1a) فإن:

$$R_t = G_t C_{t-1} G_t' \dots\dots\dots(2.4)$$

ومن المعادلة (٢,٢) فإن المصفوفة  $(C_{t-1})$  يمكن كتابتها بالصيغة الآتية:

$$C_{t-1} = C_{t-1}^{1/2} (C_{t-1}^{1/2})' \dots\dots\dots(2.5)$$

والآن من المعادلتين (٢,٤) و(٢,٥) فإن:

$$R_t = G_t C_{t-1}^{1/2} (C_{t-1}^{1/2})' G_t' = R_t^{1/2} (R_t^{1/2})'$$

حيث أن:

$$R_t^{1/2} = G_t C_{t-1}^{1/2}$$

والآن نقوم بإيجاد صيغة تعاقبية لمصفوفة ربحية كالمن وكالاتي:  
من المعادلة (2.1d):

$$K_t = R_t F_t' (F_t R_t F_t' + \sigma_\varepsilon^2)^{-1}$$

وبتحليل المصفوفة  $(R_t)$  إلى مصفوفتي الجذر ومبدلها، فإن:

$$K_t = R_t^{1/2} (R_t^{1/2})' F_t' (F_t R_t^{1/2} (R_t^{1/2})' F_t' + \sigma_\varepsilon^2)^{-1}$$

ولتكن:

$$\alpha_t = (F_t R_t^{1/2} (R_t^{1/2})' F_t' + \sigma_\varepsilon^2)^{-1}$$

حيث أن  $(\alpha)$  في حقيقته أحادي البعد (Scalar)، وليكن أيضاً:

$$H_t = (R_t^{1/2})' F_t'$$

فعدنذ:

$$\alpha_t = (H_t' H_t + \sigma_\varepsilon^2)^{-1}$$

إذن ستكون الصيغة النهائية لمصفوفة ربحية المرشح كالاتي:

$$K_t = \alpha_t R_t^{1/2} H_t \dots\dots\dots(2.6)$$

ولغرض إيجاد صيغة تعاقبية لمصفوفة التباين  $(C_t)$  نتبع ما يأتي:  
من المعادلة (2.1c):

$$\begin{aligned} C_t &= R_t - K_t F_t R_t \\ &= (I - K_t F_t) R_t \end{aligned}$$

وبتحليل المصفوفة  $(R_t)$  إلى مصفوفتي الجذر ومبدلها والتعويض بالمعادلة (٢, ٦) ، فإن:

$$\begin{aligned} C_t &= (I - \alpha_t R_t^{1/2} H_t F_t) R_t^{1/2} (R_t^{1/2})' \\ &= R_t^{1/2} (I - \alpha_t H_t H_t') (R_t^{1/2})' \end{aligned}$$

وقد رأى بوتتر أنه إذا تمكنا من إيجاد المصفوفة  $(A_t)$  التي تحقق:

$$A_t A_t' = (I - \alpha_t H_t H_t')$$

فعدنذ يمكن إيجاد الجذر التربيعي للمصفوفة  $(C_t)$  كالاتي:

$$C_t = R_t^{1/2} A_t A_t' (R_t^{1/2})' = C_t^{1/2} (C_t^{1/2})'$$

$$C_t^{1/2} = R_t^{1/2} A_t \quad \dots\dots\dots(2.7)$$

ولإيجاد المصفوفة ( $A_t$ ) نفرض أن  $\gamma_t$  تحقق الآتي:

$$A_t A_t' = (I - \alpha_t H_t H_t')$$

$$= (I - \gamma_t \alpha_t H_t H_t')(I - \gamma_t \alpha_t H_t H_t)'$$

أي أن:

$$A_t = (I - \gamma_t \alpha_t H_t H_t') \quad \dots\dots\dots(2.8)$$

وبفتح الأقواس:

$$I - \alpha_t H_t H_t' = I - 2\gamma_t \alpha_t H_t H_t' + \gamma_t^2 \alpha_t^2 H_t H_t' H_t' H_t$$

ولتكن:

$$\beta_t = H_t' H_t$$

حيث أن ( $\beta_t$ ) في حقيقته أحادي البعد (Scalar)، فعندئذٍ:

$$I - \alpha_t H_t H_t' = I - 2\gamma_t \alpha_t H_t H_t' + \gamma_t^2 \alpha_t^2 \beta_t H_t H_t'$$

ومنها:

$$(\gamma_t^2 \alpha_t^2 \beta_t - 2\gamma_t \alpha_t + \alpha_t) H_t H_t' = 0$$

$$\Rightarrow (\alpha_t \beta_t \gamma_t^2 - 2\gamma_t + 1) \alpha_t H_t H_t' = 0$$

إن ( $\alpha_t H_t H_t' = 0$ ) يعتبر حلاً غير مقبول لذلك يهمل، إذن:

$$\alpha_t \beta_t \gamma_t^2 - 2\gamma_t + 1 = 0$$

والحل بالنسبة لـ ( $\gamma_t$ ) نتوصل إلى [Maybeck, 1979]:

$$\gamma_t = \frac{1}{1 + \sqrt{\alpha_t \sigma_\varepsilon^2}}$$

وبتعويض المعادلة (٢,٨) في المعادلة (٢,٧):

$$C_t^{1/2} = R_t^{1/2} A_t$$

$$= R_t^{1/2} (I - \gamma_t \alpha_t H_t H_t')$$

وبالتعويض بالمعادلة (٢,٦) فإن:

$$C_t^{1/2} = R_t^{1/2} - \gamma_t K_t H_t' \quad \dots\dots\dots(2.9)$$

والمعادلة (٢,٩) تمثل الصيغة التعااقبية للجذر التربيعي لمصفوفة التباين ( $C_t$ ) بطريقة بوتنر.

ويمكن تلخيص خطوات خوارزمية بوتنر كالاتي:

$$H_t = (R_t^{1/2})' F_t' \quad ١. \text{ إيجاد المصفوفة } (H_t):$$

$$\alpha_t = (H_t' H_t + \sigma_\varepsilon^2)^{-1} \quad ٢. \text{ إيجاد } (\alpha_t):$$

$$K_t = \alpha_t R_t^{1/2} H_t \quad ٣. \text{ إيجاد مصفوفة ربحية المرشح } (K_t):$$

$$M_t = G_t M_{t-1} + K_t (y_t - F_t' G_t M_{t-1}) \quad ٤. \text{ إيجاد متوسط التوزيع اللاحق } (M_t):$$

$$\gamma_t = \frac{1}{1 + \sqrt{\alpha_t \sigma_\varepsilon^2}} \quad ٥. \text{ إيجاد } (\gamma_t):$$

$$C_t^{1/2} = R_t^{1/2} - \gamma_t K_t H_t' \quad ٦. \text{ إيجاد مصفوفة الجذر التربيعي } (C_t^{1/2}):$$

$$C_t = C_t^{1/2} (C_t^{1/2})' \quad ٧. \text{ إيجاد مصفوفة تباين التوزيع اللاحق } (C_t):$$

بقي أن نشير إلى أنه في خوارزمية بوتنر يمكن استخدام كلتا الطريقتين الموضحتين مسبقاً في إيجاد مصفوفة الجذر التربيعي للمصفوفة.

أنظر [Carlson, 1973; Bierman, 1977; Maybeck, 1979]

## ٢-٨-٣- مرشح كالمن-شميدت الجذر التربيعي

### Square Root Kalman-Schmidt Filter:

تم وضع هذه الخوارزمية من قبل [Bellantoni, & Dodge, 1967] حيث تتطلب هذه الخوارزمية إيجاد الجذر التربيعي لمصفوفة التباين باستخدام التحليل الطيفي للمصفوفة الذي تم توضيحه سابقاً، وتتميز هذه الخوارزمية بعدم تقيدها بالشرطين الواجب توافرها في خوارزمية بوتنر. وفيما يأتي توضيح لهذه الخوارزمية:

إذا كان لدينا النموذج الخطي الحركي المعرف بمنظومة المعادلات (١,٤)، نقوم بإيجاد مصفوفة الجذر التربيعي لمصفوفة تباين التوزيع الأولي ( $R_t$ ) باستخدام التحليل الطيفي، وكذلك لمصفوفة تباين خطأ المشاهدة ( $V_t$ ) وكالاتي:

$$\begin{aligned} R_t &= E_R \Lambda_R E_R' \\ &= [E_R \Lambda_R^{1/2} E_R'] [E_R \Lambda_R^{1/2} E_R']' \\ &= (R_t^{1/2}) (R_t^{1/2})' \end{aligned}$$

$$\begin{aligned} V_t &= E_V \Lambda_V E_V' \\ &= [E_V \Lambda_V^{1/2} E_V'] [E_V \Lambda_V^{1/2} E_V']' \\ &= (V_t^{1/2}) (V_t^{1/2})' \end{aligned}$$

وكما أثبتنا سابقاً في المعادلة (١,١٠) فإن مصفوفة تباين التوزيع اللاحق تكون بالصيغة الآتية:

$$C_t = (R_t^{-1} + F_t V_t^{-1} F_t)^{-1}$$

وبأخذ المعكوس للطرفين وتحليل المصفوفتين  $R_t$  و  $V_t$  إلى صيغة الجذر التربيعي:

$$C_t^{-1} = R_t^{-1} + F_t V_t^{-1} F_t$$

وبالضرب بالمصفوفة  $R_t$  ومعكوسها في الحد الثاني من الطرف الأيمن، وتحليلهما إلى مصفوفتي الجذر ومبدلها، وكذلك تحليل المصفوفة  $(V_t)$ ، نحصل على الآتي:

$$\begin{aligned} C_t^{-1} &= R_t^{-1} + [(R_t^{-1})^{1/2}]' (R_t^{1/2})' F_t [(V_t^{-1})^{1/2}]' (V_t^{-1})^{1/2} F_t R_t^{1/2} (R_t^{-1})^{1/2} \\ &= [(R_t^{-1})^{1/2}]' (R_t^{-1})^{1/2} + [(R_t^{-1})^{1/2}]' B B' (R_t^{-1})^{1/2} \end{aligned}$$

$$C_t^{-1} = [(R_t^{-1})^{1/2}]' (I_n + B B') (R_t^{-1})^{1/2} \dots\dots\dots (2.10)$$

حيث أن المصفوفة  $(B)$  ذات بعد  $(n \times m)$  وتساوي:

$$B = \left( (V_t^{-1})^{1/2} F_t R_t^{1/2} \right)'$$

والآن بأخذ المعكوس لطرفي المعادلة (٢، ١٠) وتحليل المصفوفة  $(B B')$  التي أبعادها  $(n \times n)$  باستخدام التحليل الطيفي:

$$\begin{aligned} C_t &= R_t^{1/2} (I_n + E_B \Lambda_B E_B')^{-1} (R_t^{1/2})' \\ &= R_t^{1/2} E_B (I_n + \Lambda_B)^{-1/2} (I_n + \Lambda_B)^{-1/2} E_B' (R_t^{1/2})' \\ &= \{ R_t^{1/2} E_B (I_n + \Lambda_B)^{-1/2} E_B' \} \times \{ E_B (I_n + \Lambda_B)^{-1/2} E_B' (R_t^{1/2})' \} \\ &= (C_t^{1/2}) (C_t^{1/2})' \end{aligned}$$

حيث أن المصفوفة  $(I_n)$  مصفوفة واحدة (Unit Matrix) ذات بعد  $(n \times n)$ ، وأن:

$$C_t^{1/2} = R_t^{1/2} E_B (I_n + \Lambda_B)^{-1/2} E_B'$$

وتمثل الصيغة الأخيرة الصيغة التعااقبية لمصفوفة الجذر التربيعي لمصفوفة التباين  $(C_t)$  حسب خوارزمية كالمن-شميدت. إلا أن هذه الصيغة لا تستخدم إلا في الحالات التي تكون فيها  $(n \leq m)$  أي عندما يكون عدد المعلمات أقل أو يساوي عدد المشاهدات في النموذج المدروس. أما في الحالات التي يكون فيها عدد المشاهدات أكثر من عدد المعلمات  $(n > m)$  فإن المصفوفة  $(\Lambda_B)$  سوف تحوي على الأقل على  $(n - m)$  من الجذور المميزة الصفيرية فنقوم بحذف هذه الجذور الصفيرية والمتجهات المميزة المقابلة لها من مصفوفة المتجهات المميزة. [Bellantoni, & Dodge, 1967]، وكالآتي:

$$\Lambda_{B_m} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m \end{bmatrix}$$

$$E_{B_1} = [e_1 \quad e_2 \quad \cdots \quad e_m]$$

حيث أن:

$\Lambda_{B_m}$ : تمثل مصفوفة القيم الذاتية غير الصفريّة، وهي ذات بُعد  $(m \times m)$ .  
 $E_{B_1}$ : مصفوفة المتجهات الذاتية المقابلة للقيم الذاتية غير الصفريّة، وهي ذات بُعد  $(n \times m)$ .

ومن ثمّ نحصل على:

$$C_t^{1/2} = R_t^{1/2} [E_{B_1} (I_m + \Lambda_{B_m})^{-1/2} E_{B_1}' + E_{B_2} E_{B_2}'] \quad \dots\dots\dots(2.11)$$

حيث أن:

$$E_{B_2} E_{B_2}' = (I_n - E_{B_1} E_{B_1}')_{(n \times n)}$$

وأن  $(I_n)$  مصفوفة واحدة ذات بعد  $(m \times m)$ .

تمثل الصيغة (٢, ١١) مصفوفة الجذر التربيعي لخوارزمية كالمن-شميدت في الحالات التي تكون

فيها  $(n > m)$ .

ولغرض إيجاد صيغة تعاقبية لحساب ربحية المرشح نتبع ما يأتي:

من المعادلة (2.1d) فإن:

$$K_t = R_t F_t' (F_t R_t F_t' + V_t)^{-1}$$

ومن تحليل المصفوفتين  $R_t$  و  $V_t$  إلى صيغة الجذر التربيعي:

$$K_t = R_t^{1/2} (R_t^{1/2})' F_t [F_t R_t^{1/2} (R_t^{1/2})' F_t' + V_t^{1/2} (V_t')^{1/2}]^{-1}$$

وبإخراج المصفوفة  $(V_t^{-1})$  كعامل مشترك وتحليل المصفوفتين  $(R_t$  و  $V_t^{-1})$  إلى صيغة الجذر

التربيعي للمصفوفة، نحصل على الآتي:

$$K_t = R_t^{1/2} (R_t^{1/2})' F_t [(V_t^{-1})']^{1/2} [(V_t^{-1})^{1/2} F_t R_t^{1/2} (R_t^{1/2})' F_t [(V_t^{1/2})^{-1}]' + I_m]^{-1} (V_t^{-1})^{1/2}$$

بذلك فإن الصيغة التعاقبية لمصفوفة ربحية المرشح تكون كالآتي:

$$K_t = R_t^{1/2} B [I_m + B' B]^{-1} (V_t^{-1})^{1/2} \quad \dots\dots\dots(2.12)$$

تمثل المعادلة (٢, ١٢) الصيغة التعااقبية لمصفوفة ربحية المرشح حسب خوارزمية كالمن-شميدت للجذر التربيعي.

ويمكن تلخيص هذه الخوارزمية بالخطوات الآتية:

$$1. \text{ إيجاد المصفوفة } (B): B = \left( (V_t^{-1})^{1/2} F_t R_t^{1/2} \right)'$$

$$2. \text{ إيجاد مصفوفة ربحية المرشح } (K_t): K_t = R_t^{1/2} B [I_m + B' B]^{-1} (V_t^{-1})^{1/2}$$

$$3. \text{ إيجاد متوسط التوزيع اللاحق } (M_t): M_t = G_t M_{t-1} + K_t (Y_t - F_t G_t M_{t-1})$$

٤. تحليل الطيف للمصفوفة  $(BB')$  وحذف الجذور المميزة الصفريّة من المصفوفة  $(\Lambda_B)$  والمتجهات المميزة المقابلة لها من المصفوفة  $(E_B)$ ، وحساب المصفوفة  $(E_{B_2} E_{B_2}')$  وذلك في الحالات التي تكون فيها  $(n > m)$ :

$$E_{B_2} E_{B_2}' = (I_n - E_{B_1} E_{B_1}')_{(n \times n)}$$

٥. إيجاد مصفوفة الجذر التربيعي  $(C_t^{1/2})$ :

$$C_t^{1/2} = R_t^{1/2} E_B (I_n + \Lambda_B)^{-1/2} E_B' \quad \text{عندما } (n \leq m)$$

$$C_t^{1/2} = R_t^{1/2} [E_{B_1} (I_m + \Lambda_{B_m})^{-1/2} E_{B_1}' + E_{B_2} E_{B_2}'] \quad \text{عندما } (n > m)$$

$$6. \text{ إيجاد مصفوفة تباين التوزيع اللاحق } (C_t): C_t = C_t^{1/2} (C_t^{1/2})'$$

إن هذه الخوارزمية فعالة، ولكنها تكون غير كفوءة في الحالات التي تكون فيها  $(n > m)$  نظراً للتعقيد الشديد في العمليات الحسابية.

أنظر [Bellantoni, & Dodge, 1967; Kaminski, et al. 1971]

## ٢-٨-٤- مرشح الجذر التربيعي لمصفوفة التباين والتباين المشترك

### Square Root Covariance Matrix Filter:

تمّ وضع هذه الخوارزمية من قبل [Andrews, 1968]، وتشرط هذه الخوارزمية إيجاد مصفوفة الجذر التربيعي باستخدام تحليل چولسكي الذي تمّ توضيحه سابقاً، وتعد هذه الخوارزمية أكفأ من خوارزمية كالمن-شميدت عندما تكون  $(m < n)$ .

وفيما يأتي توضيح لهذه الخوارزمية:

من المعادلتين (2.1c) و(2.1d):

$$C_t = R_t - R_t F_t' (F_t R_t F_t' + V_t)^{-1} F_t R_t$$



وبتحليل المصفوفة ( $R_t$ ) باستخدام تحليل چولسكي نحصل على:

$$C_t = R_t^{1/2} (R_t^{1/2})' - R_t^{1/2} (R_t^{1/2})' F_t' [F_t R_t^{1/2} (R_t^{1/2})' F_t' + V_t]^{-1} F_t R_t^{1/2} (R_t^{1/2})'$$

ولتكن:

$$Z_t = (R_t^{1/2})' F_t'$$

عندئذٍ:

$$C_t = R_t^{1/2} (R_t^{1/2})' - R_t^{1/2} Z_t [Z_t' Z_t + V_t]^{-1} Z_t' (R_t^{1/2})'$$

ولتكن ( $U_t$ ) مصفوفة الجذر التربيعي للمصفوفة ( $Z_t' Z_t + V_t$ ) حسب تحليل چولسكي، أي أن:

$$Z_t' Z_t + V_t = U_t U_t'$$

إذن:

$$C_t = R_t^{1/2} (R_t^{1/2})' - R_t^{1/2} Z_t [U_t U_t']^{-1} Z_t' (R_t^{1/2})'$$

نضرب الحد الثاني من الطرف الأيمن بالمقدار الآتي:

$$(U_t + V_t^{1/2})^{-1} [(U_t + V_t^{1/2})(U_t + V_t^{1/2})'] [(U_t + V_t^{1/2})']^{-1} = I_m$$

حيث أن المصفوفة ( $V_t^{1/2}$ ) تمثل مصفوفة الجذر التربيعي للمصفوفة ( $V_t$ ) حسب تحليل چولسكي.

$$C_t = R_t^{1/2} \{I_n - Z_t (U_t')^{-1} (U_t + V_t^{1/2})^{-1} [(U_t + V_t^{1/2})(U_t + V_t^{1/2})'] \times [(U_t + V_t^{1/2})']^{-1} U_t^{-1} Z_t'\} (R_t^{1/2})'$$

ومن ثمّ نفتح الأقواس ونجمع ونطرح بالمصفوفة ( $U_t U_t'$ ) في الجزء الآتي:

$$[(U_t + V_t^{1/2})(U_t + V_t^{1/2})'] = [U_t U_t' + U_t (V_t^{1/2})' + V_t^{1/2} U_t' + V_t^{1/2} (V_t^{1/2})' + U_t U_t' - U_t U_t'] \\ = [U_t (U_t + V_t^{1/2})' + (U_t + V_t^{1/2}) U_t' - Z_t' Z_t] \quad \dots\dots\dots(2.13)$$

وعند التعويض بالمعادلة (٢, ١٣) وفتح الأقواس ومن ثمّ التحليل باستخدام التجربة، نتوصل إلى:

$$C_t = R_t^{1/2} \{I_n - Z_t (U_t')^{-1} (U_t + V_t^{1/2})^{-1} Z_t'\} \times \{I_n - Z_t (U_t')^{-1} (U_t + V_t^{1/2})^{-1} Z_t'\} (R_t^{1/2})' \\ = C_t^{1/2} (C_t^{1/2})' \\ C_t^{1/2} = R_t^{1/2} \{I_n - Z_t (U_t')^{-1} (U_t + V_t^{1/2})^{-1} Z_t'\} \quad \dots\dots\dots(2.14)$$

حيث تمثل المعادلة (٢, ١٤) الصيغة التعاقبية لمصفوفة الجذر التربيعي لمصفوفة التباين. وجدير بالذكر أن مصفوفة الجذر هذه لا تكون بالضرورة - بشكل عام- مصفوفة مثلثية سفلى.

أما مصفوفة ربحية المرشح فيمكن أن نتوصل إلى الصيغة الخاصة بها بسهولة، حيث تساوي:

$$K_t = R_t^{1/2} Z_t (U_t U_t')^{-1} \quad \dots\dots\dots(2.15)$$

ويمكن تلخيص هذه الخوارزمية بالخطوات الآتية:

١. إيجاد مصفوفة الجذر التربيعي ( $R_t^{1/2}$ )، وإيجاد ( $Z_t$ ):  $Z_t = (R_t^{1/2})' F_t'$
٢. إيجاد مصفوفة الجذر التربيعي ( $U_t$ ):  $U_t = (Z_t' Z_t + V_t)^{1/2}$
٣. إيجاد مصفوفة ربحية المرشح ( $K_t$ ):  $K_t = R_t^{1/2} Z_t (U_t U_t')^{-1}$
٤. إيجاد متوسط التوزيع اللاحق ( $M_t$ ):  $M_t = G_t M_{t-1} + K_t (Y_t - F_t G_t M_{t-1})$
٥. إيجاد مصفوفة الجذر التربيعي ( $C_t^{1/2}$ ):  $C_t^{1/2} = R_t^{1/2} \{ I_n - Z_t (U_t U_t')^{-1} (U_t + V_t^{1/2})^{-1} Z_t' \}$
٦. إيجاد مصفوفة تباين التوزيع اللاحق ( $C_t$ ):  $C_t = C_t^{1/2} (C_t^{1/2})'$

أنظر [Andrews, 1968; Kaminski, et al. 1971; Gelb, 1974]

وهناك خوارزميات أخرى لمرشح كالمن الجذر التربيعي سوف لا نتطرق إليها بالتفصيل وإنما سنكتفي بذكرها فقط مع توضيح بسيط:

#### ١. مرشح كالمن الجذر التربيعي للمعلومات:

##### **Kalman Square Root Information Filter:**

ويعرف أيضاً بمرشح الجذر التربيعي لمعكوس مصفوفة التباين والتباين المشترك (Inverse Covariance Square Root Filter). وتعتمد هذه الخوارزمية على مصفوفة المعلومات (Information Matrix) بدلاً من مصفوفة التباين، ويكون الجذر التربيعي فيها كالاتي:

$$C_t^{-1} = C_t^{-1/2} (C_t^{-1/2})'$$

وبما أن الدقة (Precision) هي معكوس التباين ( $P_t = C_t^{-1}$ )، لذا فإنه من الممكن كتابة الصيغة أعلاه بدلالة مصفوفة الدقة (Precision Matrix) وكالاتي:

$$P_t = P_t^{1/2} (P_t^{1/2})'$$

#### ٢. مرشح تحليل مصفوفة التباين والتباين المشترك إلى العوامل (U-D):

##### **(U-D) Covariance Matrix Factorization Filter:**

يتم في هذه الخوارزمية تحليل مصفوفة التباين إلى الصيغة الآتية:

$$C_t = U_t D_t U_t'$$

حيث أن ( $U_t$ ) هي مصفوفة مثلثية عليا (Upper Triangular Matrix) ولكن عناصر القطر لها تكون واحدة (أي  $u_{ii} = 1, i = 1, 2, \dots, n$ )، أما المصفوفة ( $D_t$ ) فهي مصفوفة قطرية (Diagonal) تكون عناصر القطر فيها هي نفس عناصر القطر للمصفوفة الأصلية ( $C_t$ )، وتكون صيغة الجذر في هذه الخوارزمية كالاتي:

$$C_t^{1/2} = U_t D_t^{1/2}$$

لمزيد من المعلومات أنظر [Bierman, 1977; Maybeck, 1979].

# الفصل الثالث

الجانب التطبيقي

*Application Part*

## الفصل الثالث

### الجانب التطبيقي

#### ٣-١- مقدمة:

إن معظم تطبيقات خوارزمية مرشح كالمن الجذر التربيعي تكون في المجالات التي تتطلب دقة عالية في النتائج مثل تركيز المواد الكيماوية في الأدوية أو في مجال هندسة الاتصالات مثل الرصد الجوي والفضائي ونظام تحديد المواقع العالمي (GPS) Global Position System، ... الخ. وكما هو معلوم فإن من الصعوبة الحصول على بيانات تتبع هذه التطبيقات، لذا قمنا باستخدام بيانات المحاكاة التي تم توليدها بشكل عشوائي بالاعتماد على قيم ابتدائية Initial Values افتراضية، وتم تطبيق هذه البيانات المولدة على نماذج خطية حركية افتراضية. ولقد تم تطبيق خوارزميتين فقط من خوارزميات الجذر التربيعي الثلاث اللاتي تم توضيحن في الفصل الثاني، وهما:

- خوارزمية مرشح الجذر التربيعي لبوتر.
- خوارزمية مرشح الجذر التربيعي لمصفوفة التباين والتباين المشترك.

حيث تستخدم الخوارزمية الأولى الجذر التربيعي لمصفوفة التباين الذي يتم إيجاده من خلال تحليل (تجزئة) الطيف (Spectral Decomposition)، أما الخوارزمية الثانية فتستخدم الجذر التربيعي لمصفوفة التباين الذي يتم إيجاده من خلال تحليل جولسكي (Cholesky Decomposition). ولقد تم اختيار قيم ابتدائية مختلفة في كل حالة، وتمت أيضاً مقارنة النتائج المستحصلة مع نتائج خوارزمية مرشح كالمن الاعتيادي (Ordinary Kalman Filter) لتوضيح الفرق بين الإجراءين (الاعتيادي والجذر التربيعي).

وتم استخدام برنامج (Matlab 7.0) في توليد البيانات وفي تطبيق الخوارزميات الثلاث.

#### ٣-٢- التطبيق لمرشح الجذر التربيعي لبوتر:

نفرض أنه لدينا النموذج الخطي الحركي المعرف بـ (١،١٣)، وبما أن خوارزمية بوتتر تشترط أن معادلة النظام خالية من التشويش ( $W_t = 0$ )، لذا سيعرف النموذج من خلال الرباعي الآتي:

$$\{F, G, V, W\} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, [\sigma_\varepsilon^2], [0] \right\}$$

والآن لغرض توليد البيانات لا بد من تحديد قيم ابتدائية عند الزمن ( $t = 0$ ) لكل من تباين خطأ المشاهدات  $\sigma_\varepsilon^2$  ومصفوفة تباين التوزيع الابتدائي  $C_t$  ومتجه متوسط التوزيع الابتدائي  $M_t$ ، وحسب الصيغة الآتية:

$$C_t = \begin{bmatrix} c_{11(t)} & c_{12(t)} \\ c_{21(t)} & c_{22(t)} \end{bmatrix}, \quad M_t = \begin{bmatrix} m_t \\ b_t \end{bmatrix}$$

حيث نستخدم قيمة تباين خطأ المشاهدات  $\sigma_\varepsilon^2$  في توليد حد الخطأ العشوائي  $\varepsilon_t$ ، حيث أن:

$$\varepsilon_t \stackrel{i.i.d}{\sim} N(0, \sigma_\varepsilon^2)$$

ومن ثم نعوض في معادلات النموذج لإيجاد قيم  $y_t$  لكل مرحلة زمنية. وكما أشرنا سابقاً في الفصل الثاني فإننا سنستخدم المصفوفة  $C_0$  والمتجه  $M_0$  كقيم ابتدائية في تطبيق خوارزمية مرشح الجذر التربيعي لبوتر وكذلك في تطبيق خوارزمية مرشح كالمن الاعتيادي.

ولقد تم اختيار أربع قيم ابتدائية مختلفة لمصفوفة التباين  $C_t$  وتباين خطأ المشاهدة  $\sigma_\varepsilon^2$ ، أما متجه المتوسط  $M_t$  فقد اختيرت له قيمة ابتدائية واحدة لجميع الحالات وهي:

$$M_0 = \begin{bmatrix} 6 \\ 0.5 \end{bmatrix}$$

وتم إدراج البيانات التي تم توليدها والنتائج في جداول مختصرة - في بعض الحالات - في الملحق C.

والفقرات الآتية توضح عملية الترشيح باستخدام خوارزمية بوتر، فضلاً عن مقارنة النتائج مع نتائج خوارزمية مرشح كالمن الاعتيادي في الحالات الأربع:

### ❖ الحالة الأولى:

نفرض أن القيم الابتدائية عند الزمن ( $t = 0$ ) لكل من مصفوفة التباين  $C_t$  وتباين خطأ

المشاهدة  $\sigma_\varepsilon^2$  - قيمة تباين خطأ المشاهدة ثابتة خلال الزمن - تكون كالاتي:

$$C_0 = \begin{bmatrix} 0.65 & 0 \\ 0 & 0.15 \end{bmatrix}, \quad \sigma_\varepsilon^2 = 0.25$$

وبتطبيق خوارزمية مرشح الجذر التربيعي لبوتر على هذه القيم تكون نتائج الترشيح كالاتي:

تتقارب (Convergent) مصفوفة التباين  $C_t$  عند الزمن ( $t = 38$ )، حيث تكون:

$$C_t = \begin{bmatrix} 0.0322 & 0.0022 \\ 0.0022 & 0.0003 \end{bmatrix}, \quad t = 38, 39, 40, \dots$$

ويحصل الاستقرار في ربحية المرشح  $K_t$  عند الزمن ( $t = 39$ )، حيث تكون:

$$K_t = \begin{bmatrix} 0.1286 \\ 0.0087 \end{bmatrix}, \quad t = 39, 40, 41, \dots$$

أنظر الجدول (C-1).

بينما عند تطبيق خوارزمية مرشح كالمن الاعتيادي نلاحظ عدم حصول تقارب في مصفوفة التباين، حيث تبدأ أولى المشاكل بالظهور عند المرحلة الزمنية ( $t = 39$ )، حيث تفقد هذه المصفوفة خاصية أكيدة الإيجابية (Positive Definite)، حيث تكون:

$$C_{39} = \begin{bmatrix} 0.0244 & 0.0009 \\ 0.0009 & 0 \end{bmatrix}$$

وعند الزمن ( $t = 173$ ) تتحول مصفوفة التباين إلى مصفوفة شاذة (Singular)، حيث تكون:

$$C_{173} = \begin{bmatrix} 0.0057 & 0 \\ 0 & 0 \end{bmatrix}$$

ويلاحظ أن عدم حصول تقارب في مصفوفة التباين ينعكس على ربحية المرشح، حيث لا يحصل الاستقرار فيها أبداً. أنظر الجدول (C-2).

#### ❖ الحالة الثانية:

نفرض أن القيم الابتدائية عند الزمن ( $t = 0$ ) لكل من مصفوفة التباين  $C_t$  وتباين خطأ

المشاهدة  $\sigma_\varepsilon^2$  - قيمة تباين خطأ المشاهدة ثابتة خلال الزمن- تكون كالاتي:

$$C_0 = \begin{bmatrix} 1.15 & 0 \\ 0 & 0.84 \end{bmatrix}, \quad \sigma_\varepsilon^2 = 0.5$$

وبتطبيق خوارزمية مرشح الجذر التربيعي ليوتر على هذه القيم تكون نتائج الترشيح كالاتي:

تتقارب مصفوفة التباين  $C_t$  عند الزمن ( $t = 48$ )، حيث تكون:

$$C_t = \begin{bmatrix} 0.0516 & 0.0028 \\ 0.0028 & 0.0003 \end{bmatrix}, \quad t = 48, 49, 50, \dots$$

ويحصل الاستقرار في ربحية المرشح  $K_t$  عند الزمن ( $t = 49$ )، حيث تكون:

$$K_t = \begin{bmatrix} 0.1031 \\ 0.0055 \end{bmatrix}, \quad t = 49, 50, 51, \dots$$

أنظر الجدول (C-3).

من جهة أخرى، تفشل خوارزمية مرشح كالمن الاعتيادي في الوصول إلى حالة الاستقرار (Steady State)، وذلك بسبب حصول تباعد (Divergence) في مصفوفة التباين  $C_t$ ، ويلاحظ انخفاض قيم التباينات بشكل كبير، الأمر الذي يجعل من مصفوفة التباين مصفوفة غير أكيدة الإيجابية عند الزمن ( $t = 49$ )، وتتطور المشاكل في مصفوفة التباين لتصبح مصفوفة شاذة عند الزمن ( $t = 245$ )، حيث تكون كالاتي:

$$C_{49} = \begin{bmatrix} 0.0392 & 0.0012 \\ 0.0012 & 0 \end{bmatrix}$$

$$C_{245} = \begin{bmatrix} 0.0081 & 0 \\ 0 & 0 \end{bmatrix}$$

أنظر الجدول (C-4).

#### ❖ الحالة الثالثة:

نفرض أن القيم الابتدائية عند الزمن ( $t = 0$ ) لكل من مصفوفة التباين  $C_t$  وتباين خطأ المشاهدة  $\sigma_\varepsilon^2$  - قيمة تباين خطأ المشاهدة ثابتة خلال الزمن- تكون كالاتي:

$$C_0 = \begin{bmatrix} 0.9 & 0 \\ 0 & 1.15 \end{bmatrix}, \quad \sigma_\varepsilon^2 = 0.4$$

وبتطبيق خوارزمية مرشح الجذر التربيعي ليوتر على هذه القيم تكون نتائج الترشيح كالاتي: يحصل التقارب (Convergence) في مصفوفة التباين  $C_t$  عند الزمن ( $t = 44$ )، حيث تكون:

$$C_t = \begin{bmatrix} 0.0444 & 0.0026 \\ 0.0026 & 0.0003 \end{bmatrix}, \quad t = 44, 45, 46, \dots$$

ويحصل الاستقرار في ربحية المرشح  $K_t$  عند الزمن ( $t = 45$ )، حيث تكون:

$$K_t = \begin{bmatrix} 0.1109 \\ 0.0064 \end{bmatrix}, \quad t = 45, 46, 47, \dots$$

أنظر الجدول (C-5).

وعند تطبيق خوارزمية مرشح كالمن الاعتيادي على القيم الابتدائية المختارة، فإن تفوق خوارزمية بوتر للترشيح يظهر بشكل واضح، حيث تظهر المشاكل في مصفوفة التباين  $C_t$  عند الزمن ( $t = 46$ )، حيث تكون:

$$C_t = \begin{bmatrix} 0.0333 & 0.0011 \\ 0.0011 & 0 \end{bmatrix}$$

أي تكون غير أكيدة الإيجابية، وعند الزمن ( $t = 219$ ) تتعقد المشاكل في مصفوفة التباين وتتحول إلى مصفوفة شاذة، حيث تكون:

$$C_t = \begin{bmatrix} 0.0072 & 0 \\ 0 & 0 \end{bmatrix}$$

وهذه المشاكل تنعكس بالتأكيد على استقرارية الخوارزمية، حيث لا يحصل الاستقرار نهائياً في ربحية المرشح  $K_t$ ، وبذلك تحدث مشكلة التباعد (Divergence Problem) في الخوارزمية. أنظر الجدول (C-6)، لذا تكون نتائج الترشيح غير مقبولة.

#### ❖ الحالة الرابعة:

نفرض أن القيم الابتدائية عند الزمن ( $t=0$ ) لمصفوفة التباين  $C_t$  وتباين خطأ المشاهدة  $\sigma_\varepsilon^2$  - قيمة تباين خطأ المشاهدة ثابتة خلال الزمن- تكون كالآتي:

$$C_0 = \begin{bmatrix} 1.6 & 0 \\ 0 & 1.8 \end{bmatrix}, \quad \sigma_\varepsilon^2 = 0.35$$

وكما في الحالات السابقة أعلاه، فإن تطبيق خوارزمية الجذر التربيعي ليوتر تجعل مصفوفة التباين تتقارب، وبالتالي تحدث حالة الاستقرار في ربحية المرشح. حيث يحدث التقارب في مصفوفة التباين  $C_t$  عند الزمن ( $t = 41$ )، فيما يحدث الاستقرار ربحية المرشح  $K_t$  في المرحلة ( $t = 42$ )، وتكون القيم كالآتي:

$$C_t = \begin{bmatrix} 0.0400 & 0.0024 \\ 0.0024 & 0.0003 \end{bmatrix}, \quad t = 41, 42, 43, \dots$$

$$K_t = \begin{bmatrix} 0.1141 \\ 0.0068 \end{bmatrix}, \quad t = 42, 43, 44, \dots$$

أنظر الجدول (C-7).

بالمقابل فإن تطبيق خوارزمية مرشح كالمن الاعتيادي تفشل في تحقيق الاستقرار، وذلك يعود لعدم تقارب مصفوفة التباين  $C_t$ . فيما تظهر المشاكل نفسها التي واجهتنا في الحالات السابقة في



مصفوفة التباين، حيث تفقد خاصية أكيدة الإيجابية عند الزمن ( $t = 44$ ) و تصبح مصفوفة شاذة عند الزمن ( $t = 205$ )، حيث تكون:

$$C_{44} = \begin{bmatrix} 0.0306 & 0.0010 \\ 0.0010 & 0 \end{bmatrix}$$

$$C_{205} = \begin{bmatrix} 0.0068 & 0 \\ 0 & 0 \end{bmatrix}$$

أنظر الجدول (C-8).

#### ◆ ملاحظة:

من مقارنة نتائج خوارزمية مرشح الجذر التربيعي لبوتر مع نتائج خوارزمية مرشح كالمن الاعتيادي في الحالات الأربع (أنظر الملحق C)، نلاحظ أن قيم عناصر مصفوفة التباين  $C_t$  ومصفوفة ربحية المرشح  $K_t$  تكون متساوية في كلتا الخوارزميتين - الاعتيادية والجذر التربيعي- في المراحل الزمنية الأولى من عملية الترشيح، وذلك لأن المشاكل العددية لا تظهر في وقت مبكر، وإنما تظهر نتيجة للزيادة المؤثرة في عدد الإجراءات التعاقبية. فمثلاً في الحالة الأولى تكون مصفوفة التباين وربحية المرشح في الزمن ( $t = 1$ ) في خوارزمية مرشح الجذر التربيعي لبوتر كالآتي:

$$C_1 = \begin{bmatrix} 0.1905 & 0.0357 \\ 0.0357 & 0.1268 \end{bmatrix}$$

$$K_1 = \begin{bmatrix} 0.7619 \\ 0.1428 \end{bmatrix}$$

وهي القيم نفسها التي أظهرتها خوارزمية مرشح كالمن الاعتيادي، الأمر الذي يثبت أن الإجراءات يعطيان نتائج متساوية في الحالات التي لا توجد فيها مشاكل عددية. وتبقى هذه القيم متساوية حتى المرحلة الزمنية ( $t = 9$ )، حيث تكون القيم كالآتي:

$$C_9 = \begin{bmatrix} 0.0890 & 0.0148 \\ 0.0148 & 0.0035 \end{bmatrix}$$

$$K_9 = \begin{bmatrix} 0.3558 \\ 0.0592 \end{bmatrix}$$

ويلاحظ أيضاً أن هذه القيم تكون متقاربة في المراحل اللاحقة، مع ظهور بعض الاختلافات بشكل تدريجي، وتكون هذه الاختلافات واضحة عند الزمن ( $t = 23$ )، حيث تكون القيم في خوارزمية مرشح الجذر التربيعي لبوتر كالاتي:

$$C_{23} = \begin{bmatrix} 0.0412 & 0.0028 \\ 0.0028 & 0.0003 \end{bmatrix}$$

$$K_{23} = \begin{bmatrix} 0.1650 \\ 0.0113 \end{bmatrix}$$

بينما تكون القيم في خوارزمية مرشح كالمن الاعتيادي عند الزمن نفسه كالاتي:

$$C_{23} = \begin{bmatrix} 0.0400 & 0.0026 \\ 0.0026 & 0.0002 \end{bmatrix}$$

$$K_{23} = \begin{bmatrix} 0.1601 \\ 0.0104 \end{bmatrix}$$

حيث يلاحظ أن قيم عناصر مصفوفة التباين التي أظهرتها خوارزمية مرشح كالمن الاعتيادي تكون أقل من القيم التي أظهرتها خوارزمية مرشح الجذر التربيعي لبوتر، وهذا يعني أن خوارزمية مرشح كالمن الاعتيادي تعطي قيماً أصغر للتباين بشكل غير منطقي. ويلاحظ أيضاً في الحالات الأربع أن الاستقرار الكلي في ربحية المرشح يحدث في زمن واحد (أي استقرار عناصر متجه ربحية المرشح  $K_t$ )، وذلك بعد تقارب جميع عناصر مصفوفة التباين  $C_t$ ، لأن عناصر متجه ربحية المرشح تعتمد في حسابها على عناصر مصفوفة تباين التوزيع السابق  $R_t$ ، وتباين خطأ المشاهدة  $\sigma_\varepsilon^2$ ، حيث تكون كالاتي:

$$K_t = \begin{bmatrix} k_{11(t)} \\ k_{21(t)} \end{bmatrix} = \begin{bmatrix} r_{11(t)} / (r_{11(t)} + \sigma_\varepsilon^2) \\ r_{12(t)} / (r_{11(t)} + \sigma_\varepsilon^2) \end{bmatrix}$$

وبما أن  $\sigma_\varepsilon^2$  ثابتة في جميع المراحل الزمنية، فإن استقرارية  $K_t$  تعتمد على تقارب المصفوفة  $R_t$  فقط.

ومن (١٤، ١) فإن:

$$r_{11(t)} = c_{11(t-1)} + 2c_{12(t-1)} + c_{22(t-1)} + \sigma_{\delta_\mu}^2 + \sigma_{\delta_\beta}^2$$

مع ملاحظة  $\sigma_{\delta_\mu}^2 = \sigma_{\delta_\beta}^2 = 0$ ، ولأن عنصري متجه ربحية المرشح يعتمدان على العنصر  $r_{11(t)}$ ، الذي يعتمد بدوره على جميع عناصر مصفوفة التباين  $C_t$ ، لذلك فإن الاستقرار الكلي في متجه ربحية المرشح يحصل في الزمن نفسه على الرغم من استقرار عناصر مصفوفة التباين  $C_t$  في أزمنة متفاوتة. أنظر الملحق C.

### ٣-٣- التطبيق لمرشح الجذر التربيعي لمصفوفة التباين والتباين المشترك:

نفرض أنه لدينا النموذج الخطي الحركي الآتي:

$$Y_t = F_t M_t + E_t$$

$$M_t = G_t M_{t-1} + \Delta_t$$

حيث أن:

$$Y_t = \begin{bmatrix} y_{11(t)} \\ y_{21(t)} \\ y_{31(t)} \end{bmatrix}, \quad M_t = \begin{bmatrix} \mu_{11(t)} \\ \mu_{21(t)} \\ \mu_{31(t)} \end{bmatrix}, \quad E_t = \begin{bmatrix} \varepsilon_{11(t)} \\ \varepsilon_{21(t)} \\ \varepsilon_{31(t)} \end{bmatrix}, \quad \Delta = \begin{bmatrix} \delta_{11(t)} \\ \delta_{21(t)} \\ \delta_{31(t)} \end{bmatrix}$$

ويمكن وصف هذا النموذج من خلال الرباعي الآتي:

$$\{F, G, V, W\} = \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} \sigma_{\varepsilon_1}^2 & 0 & 0 \\ 0 & \sigma_{\varepsilon_2}^2 & 0 \\ 0 & 0 & \sigma_{\varepsilon_3}^2 \end{bmatrix}, \begin{bmatrix} \sigma_{\mu_1}^2 & 0 & 0 \\ 0 & \sigma_{\mu_2}^2 & 0 \\ 0 & 0 & \sigma_{\mu_3}^2 \end{bmatrix} \right\}$$

وكما في الفقرة السابقة، لغرض توليد البيانات نقوم بتحديد قيم ابتدائية عند الزمن ( $t=0$ )

لكل من مصفوفة تباين خطأ المشاهدات  $V$  ومصفوفة تباين خطأ النظام  $W$  ومصفوفة التباين  $C_t$

ومتجه المتوسط  $M_t$ ، وحسب الصيغة الآتية:

$$V = \begin{bmatrix} \sigma_{\varepsilon_1}^2 & 0 & 0 \\ 0 & \sigma_{\varepsilon_2}^2 & 0 \\ 0 & 0 & \sigma_{\varepsilon_3}^2 \end{bmatrix}, \quad W = \begin{bmatrix} \sigma_{\mu_1}^2 & 0 & 0 \\ 0 & \sigma_{\mu_2}^2 & 0 \\ 0 & 0 & \sigma_{\mu_3}^2 \end{bmatrix}$$

$$C_t = \begin{bmatrix} c_{11(t)} & 0 & 0 \\ 0 & c_{22(t)} & 0 \\ 0 & 0 & c_{33(t)} \end{bmatrix}, \quad M_t = \begin{bmatrix} m_{11(t)} \\ m_{21(t)} \\ m_{31(t)} \end{bmatrix}$$

حيث نستخدم مصفوفة تباين خطأ المشاهدات  $V$  ومصفوفة تباين خطأ النظام  $W$  في توليد

متجهي الخطأ العشوائي  $E_t$  و  $\Delta_t$  على التوالي، حيث أن:

$$E_t \stackrel{i.i.d}{\sim} N_3(0, V)$$

$$\Delta_t \stackrel{i.i.d}{\sim} N_3(0, W)$$

ومن ثم نعوض في النموذج المعرف أعلاه لإيجاد قيم  $Y_t$  لكل مرحلة زمنية، وكذلك نستخدم

المصفوفة  $C_0$  والمتجه  $M_0$  في تطبيق خوارزمية مرشح الجذر التربيعي لمصفوفة التباين والتباين

المشترك وكذلك في تطبيق خوارزمية مرشح كالمن الاعتيادي.

ولقد تمّ اختيار ثلاثة قيم ابتدائية مختلفة لمصفوفة التباين  $C_t$  ومصفوفة تباين خطأ المشاهدات  $V$  ومصفوفة تباين خطأ النظام  $W$ ، أما متجه المتوسط  $M_t$  فقد اختيرت له قيمة ابتدائية واحدة لجميع الحالات وهي:

$$M_0 = \begin{bmatrix} 10 \\ 2.4 \\ 0.6 \end{bmatrix}$$

وكذلك أدرجت البيانات التي ولدت والنتائج في جداول مختصرة - في بعض الحالات- في الملحق D. والفقرات الآتية توضح عملية الترشيح باستخدام خوارزمية الجذر التربيعي لمصفوفة التباين والتباين المشترك، وكذلك مقارنة النتائج مع نتائج خوارزمية مرشح كالمن الاعتيادي في الحالات الثلاث:

#### ❖ الحالة الأولى:

نفرض أن القيم الابتدائية عند الزمن ( $t = 0$ ) لكل من مصفوفة التباين  $C_t$  ومصفوفة تباين خطأ المشاهدات  $V$  ومصفوفة تباين خطأ النظام  $W$  تكون كالاتي:

$$C_0 = \begin{bmatrix} 0.7 & 0 & 0 \\ 0 & 0.09 & 0 \\ 0 & 0 & 0.03 \end{bmatrix}, V = \begin{bmatrix} 0.015 & 0 & 0 \\ 0 & 0.04 & 0 \\ 0 & 0 & 0.063 \end{bmatrix}$$

$$W = \begin{bmatrix} 0.0000001 & 0 & 0 \\ 0 & 0.0000004 & 0 \\ 0 & 0 & 0.0000005 \end{bmatrix}$$

حيث كانت نتائج الترشيح باستخدام خوارزمية الجذر التربيعي كالاتي:  
يحصل التقارب في مصفوفة التباين  $C_t$  في الزمن ( $t = 30$ )، حيث تكون:

$$C_t = \begin{bmatrix} 0.0008 & 0 & 0 \\ 0 & 0.0014 & 0 \\ 0 & 0 & 0.0018 \end{bmatrix}, \quad t = 30, 31, 32, \dots$$

ويحصل الاستقرار في الخوارزمية باستقرار ربحية المرشح  $K_t$  بشكل كلي عند الزمن

( $t = 31$ )، حيث تكون:

$$K_t = \begin{bmatrix} 0.0506 & 0 & 0 \\ 0 & 0.0338 & 0 \\ 0 & 0 & 0.0277 \end{bmatrix}, \quad t = 31,32,33,\dots$$

أنظر الجدول (D-1).

من جهة أخرى، عند تطبيق القيم الابتدائية نفسها على خوارزمية مرشح كالمن الاعتيادي، فإن الخوارزمية تفشل في الوصول إلى حالة الاستقرار، فبالرغم من ثبات مصفوفة التباين على قيم محددة لفترة زمنية معينة إلا أن الاستقرار في ربحية المرشح لا يحصل أبداً، وذلك يجعلنا نستمر بعملية الترشيح لفترات زمنية طويلة.

ونظراً لزيادة التكرارات فإن المشاكل في مصفوفة التباين لا بُدَّ وأن تظهر، ومن ملاحظة النتائج فإن مصفوفة التباين تتحول إلى مصفوفة غير أكيدة الإيجابية وشاذة في الوقت نفسه، وذلك عند الزمن ( $t = 399$ )، وكالاتي:

$$C_{399} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.0001 & 0 \\ 0 & 0 & 0.0002 \end{bmatrix}$$

ويلاحظ أن قيم عناصر مصفوفة التباين تكون متساوية (أو متقاربة) في المراحل الزمنية الابتدائية في نتائج كلتا الخوارزميتين، حيث تكون عند الزمن ( $t = 7$ ) متساوية تماماً وتكون القيم كالاتي:

$$C_7 = \begin{bmatrix} 0.0021 & 0 & 0 \\ 0 & 0.0054 & 0 \\ 0 & 0 & 0.0069 \end{bmatrix}$$

وتبقى هذه القيم متقاربة لعدد من الفترات الزمنية اللاحقة، وهذا يدل على قلة المشاكل العددية في بداية عملية الترشيح. أنظر الجدول (D-2).

### ❖ الحالة الثانية:

نفرض أن القيم الابتدائية عند الزمن ( $t = 0$ ) لكل من مصفوفة التباين  $C_t$  ومصفوفة تباين خطأ المشاهدات  $V$  ومصفوفة تباين خطأ المعلمات  $W$  تكون كالاتي:

$$C_0 = \begin{bmatrix} 0.153 & 0 & 0 \\ 0 & 0.051 & 0 \\ 0 & 0 & 1.19 \end{bmatrix}, \quad V = \begin{bmatrix} 0.036 & 0 & 0 \\ 0 & 0.0576 & 0 \\ 0 & 0 & 0.0135 \end{bmatrix}$$

$$W = \begin{bmatrix} 0.00000052 & 0 & 0 \\ 0 & 0.00000065 & 0 \\ 0 & 0 & 0.00000013 \end{bmatrix}$$

و كانت نتائج الترشيح باستخدام خوارزمية الجذر التربيعي كالآتي:

يحصل التقارب في مصفوفة التباين  $C_t$  في الزمن ( $t = 31$ )، حيث تكون:

$$C_t = \begin{bmatrix} 0.0013 & 0 & 0 \\ 0 & 0.0017 & 0 \\ 0 & 0 & 0.0008 \end{bmatrix}, \quad t = 31, 32, 33, \dots$$

وتستقر الخوارزمية عند الزمن ( $t = 32$ ) باستقرار ربحية المرشح بشكل كلي، حيث تكون

كالآتي:

$$K_t = \begin{bmatrix} 0.0348 & 0 & 0 \\ 0 & 0.0286 & 0 \\ 0 & 0 & 0.0559 \end{bmatrix}, \quad t = 32, 33, 34, \dots$$

أنظر الجدول (D-3).

أما عند تطبيق خوارزمية مرشح كالمن الاعتيادي على القيم الابتدائية أعلاه، فإننا نلاحظ عدم تقارب مصفوفة التباين إلى قيمة معينة، حيث تستمر عناصر مصفوفة التباين بالانخفاض التدريجي حتى تتحول إلى مصفوفة شاذة في المرحلة ( $t = 391$ )، حيث تكون كالآتي:

$$C_{391} = \begin{bmatrix} 0.0002 & 0 & 0 \\ 0 & 0.0002 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

وهذا يؤثر بشكل سلبي على استقرار الخوارزمية، حيث أن الاستقرار لا يحدث أبداً في مصفوفة ربحية المرشح، وبالتالي يفشل الإجراء، وتحدث حالة التباعد، لذا لا يمكن الاعتماد على نتائج الترشيح. أنظر الجدول (D-4).

❖ الحالة الثالثة:

نفرض أن القيم الابتدائية عند الزمن ( $t = 0$ ) لكل من مصفوفة التباين  $C_t$  ومصفوفة تباين

خطأ المشاهدات  $V$  ومصفوفة تباين خطأ النظام  $W$  تكون كالآتي:

$$C_0 = \begin{bmatrix} 0.0321 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 0.0107 \end{bmatrix}, \quad V = \begin{bmatrix} 0.0216 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.0345 \end{bmatrix}$$

$$W = \begin{bmatrix} 0.00000078 & 0 & 0 \\ 0 & 0.000000195 & 0 \\ 0 & 0 & 0.000000975 \end{bmatrix}$$

وبتطبيق خوارزمية الجذر التربيعي لمصفوفة التباين والتباين المشترك فإن نتائج الترشيح تكون كالآتي:

يحصل التقارب في مصفوفة التباين عند الزمن ( $t=21$ )، حيث تتقارب عناصر مصفوفة التباين إلى القيم الآتية:

$$C_t = \begin{bmatrix} 0.001 & 0 & 0 \\ 0 & 0.0007 & 0 \\ 0 & 0 & 0.0013 \end{bmatrix}, \quad t = 21, 22, 23, \dots$$

ويحصل الاستقرار عند الزمن ( $t = 22$ )، حيث يكون الشكل النهائي لمصفوفة ربحية المرشح كالآتي:

$$K_t = \begin{bmatrix} 0.0442 & 0 & 0 \\ 0 & 0.0654 & 0 \\ 0 & 0 & 0.0363 \end{bmatrix}, \quad t = 22, 23, 24, \dots$$

أنظر الجدول (D-5).

وعند تطبيق خوارزمية مرشح كالمن الاعتيادي على القيم الابتدائية المختارة، فإن تفوق خوارزمية مرشح الجذر التربيعي يظهر بشكل واضح، حيث تفشل خوارزمية مرشح كالمن الاعتيادي في الوصول إلى نتائج منطقية، حيث ينخفض التباين بشكل دائم ولا يحدث الاستقرار أبداً، وتزداد المشاكل تعقيداً بتحول مصفوفة التباين إلى مصفوفة شاذة عند الزمن ( $t=313$ )، حيث تكون كالآتي:

$$C_{313} = \begin{bmatrix} 0.0001 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.0002 \end{bmatrix}$$

مما يؤدي إلى فشل الإجراء وحدوث مشكلة التباعد. أنظر الجدول (D-6).

#### ◆ ملاحظة:

بالنظر إلى نتائج الترشيح باستخدام خوارزمية مرشح الجذر التربيعي لمصفوفة التباين والتباين المشترك، نلاحظ أن الاستقرار في عناصر مصفوفة ربحية المرشح  $K_t$  يحدث في أزمنة

متفاوتة، حيث يستقر كل عنصر مباشرة بعد تقارب العنصر المقابل له في مصفوفة التباين  $C_t$ ، وذلك لاعتماد عناصر ربحية المرشح فقط على العناصر المقابلة لها في مصفوفة التباين  $C_t$ . حيث أن:

$$K_t = \begin{bmatrix} k_{11(t)} & 0 & 0 \\ 0 & k_{22(t)} & 0 \\ 0 & 0 & k_{33(t)} \end{bmatrix}$$

$$= \begin{bmatrix} r_{11(t)}/(r_{11(t)} + \sigma_{\varepsilon_1}^2) & 0 & 0 \\ 0 & r_{22(t)}/(r_{22(t)} + \sigma_{\varepsilon_2}^2) & 0 \\ 0 & 0 & r_{33(t)}/(r_{33(t)} + \sigma_{\varepsilon_3}^2) \end{bmatrix}$$

وحيث أن  $R_t$  يمكن إيجادها كالآتي:

$$R_t = \begin{bmatrix} r_{11(t)} & 0 & 0 \\ 0 & r_{22(t)} & 0 \\ 0 & 0 & r_{33(t)} \end{bmatrix}$$

$$= \begin{bmatrix} c_{11(t-1)} + \sigma_{\delta_1}^2 & 0 & 0 \\ 0 & c_{22(t-1)} + \sigma_{\delta_2}^2 & 0 \\ 0 & 0 & c_{33(t-1)} + \sigma_{\delta_3}^2 \end{bmatrix}$$

وكما نرى فإن كل عنصر من عناصر المصفوفة  $K_t$  يعتمد فقط على العنصر المقابل له من مصفوفة التباين  $C_t$ ، الأمر الذي يؤدي إلى استقرار عناصر المصفوفة  $K_t$  في أوقات زمنية متفاوتة.



## الاستنتاجات والتوصيات

### ١. الاستنتاجات:

- مما تقدم في الجانبين النظري والتطبيقي، نستنتج ما يأتي:
- إن خوارزمية مرشح كالمن الاعتيادي لا تلائم جميع أنواع البيانات.
- إن استخدام خوارزمية مرشح كالمن الجذر التربيعي في الدراسات التطبيقية التي تتطلب دقة عالية في النتائج يكون أفضل من استخدام خوارزمية مرشح كالمن الاعتيادي، حيث تقودنا الأولى إلى الاستقرار بينما تقودنا الثانية إلى حدوث مشكلة التباعد (Divergence)، وفشل الإجراء.
- إن استخدام خوارزمية مرشح كالمن الجذر التربيعي يعطي نتائج متطابقة مع نتائج خوارزمية مرشح كالمن الاعتيادي في الحالات التي تقل فيها المشاكل العددية، أو عندما يكون عدد التكرارات قليل نسبياً، كما لوحظ في الجانب العملي، وهذا يدل على عدم اختلاف الإجراءين جبرياً.

### ٢. التوصيات:

- بالنظر لما تقدم نوصي بما يأتي:
- الاستفادة من الدراسة في المجالات التطبيقية المتاحة، خصوصاً في مجال الصناعات الدوائية.
- القيام بدراسة مرشح كالمن الجذر التربيعي في الحالات التي تكون فيها الأخطاء مترابطة، أو عندما تكون النماذج غير خطية مثل مرشح كالمن الجذر التربيعي الموسع (Extended Square Root Filter)، أو مرشح كالمن الجذر التربيعي للمعلومات (Square Root Information Kalman Filter).
- الاستفادة من نظام ماتلاب (Matlab) في إجراء وتطبيق خوارزميات الترشيح المتنوعة نظراً للسلاسة والدقة العاليتين اللتين يوفرهما هذا النظام.

## المصادر:

١. الحمداني، مهند سعد الله داود. (١٩٩٦)، "مرشح كالمن لبعض النماذج الحركية الخطية مع المحاكاة" رسالة ماجستير غير منشورة، كلية الإدارة والاقتصاد، جامعة الموصل.
٢. الرسام، رياسالم محمد علي. (٢٠٠٤)، "الترشيح والتنبؤ لمعلمات النماذج الحركية في الأزمنة المتقطعة باستخدام أسلوب بيز"، أطروحة دكتوراه غير منشورة، كلية علوم الحاسبات والرياضيات، جامعة الموصل.
3. Adams, R.A. (1999), "*Calculus*", Addison Wesley Longman Ltd.
4. Anderson, B. & Moore, J. (1979), "*Optimal Filtering*", Prentice Hall.
5. Andrews, A. (1968), "A Square Root Formulation of the Kalman Covariance Equations", ", *AIAA Journal*, Vol. 6, No. 6, PP. 1165-1166.
6. Barnett, V. (1982), "*Comparative Statistical Inference* ", John Wiley & Sons, New York.
7. Bellantoni, J.F. & Dodge, K.W. (1967), "A Square Root Formulation of the Kalman-Schmidt Filter", *AIAA Journal*, Vol. 5, No. 7, PP. 1309-1314.
8. Bierman, G.J. (1977), "*Factorization Methods for Discrete Sequential Estimation*", Academic Press, Inc., New York.
9. Box, G.E. & Tiao, G.C. (1973), "*Bayesian Inference in Statistical Analysis*", Addison-Wesley Publishing Company, Inc.
10. Bozic, S.M. (1979), "*Digital and Kalman Filtering*", Edward Arnold (Publishers) Ltd., London.
11. Calder, C. *et al.* (2003), "Incorporating Multiple Sources of Stochasticity Into Dynamic Population Models", *Ecology*, Vol. 84, No. 6, PP. 1395-1402.
12. Carlson, N.A. (1973), "Fast Triangular Formulation of the Square Root Filter", ", *AIAA Journal*, Vol. 11, No. 9, PP. 1259-1265.
13. Dulger, I. (2001), "**Multiple Model Adaptive Estimation for Time Series Analysis**" Unpublished M. Sc. Thesis, Dept. of the Air Force, Air University, Ohio.
14. Gelb, A. (1974), "*Applied Optimal Estimation*", The MIT Press.

15. Harrison, P.J. & Stevens, C.F. (1976), "Bayesian Forecasting", *JRSSB*, Vol.38, PP. 205-246.
16. Jalil, T.S. (1988), "**Sequential Inspection & Intervention Policy for a Manufacturing Process**", Unpublished Ph.D. Dissertation, Dept. of Mathematics, Wales University.
17. Kailath, T. (1974), "A View of Three Decades of Linear Filtering Theory", *IEEE Trans. Inform. Theory*, Vol. IT-20, No. 2, PP. 146-181.
18. Kalman, R.E. (1960), "A New Approach to Linear Filtering and Prediction Problems", *Journal of Basic Engineering*, Vol.82, PP. 35-45.
19. Kalman, R.E. & Bucy, R.S. (1961), "New Results in Linear Filtering and Prediction Theory", *Journal of Basic Engineering*, Vol.83, PP.95-108.
20. Kaminski, P.G. *et al.* (1971), "Discrete Square Root Filtering: A Survey of Current Techniques", *IEEE Trans. Automatic Control*, Vol. AC-16, No. 6, PP. 727-736.
21. Kendall, M.G. & Stuart A. (1976), "*The Advanced Theory of Statistics*", Charles Griffin & Company Limited, London. Vol. 2.
22. Lindgren, B.W. (1962), "*Statistical Theory*", Collier Macmillan, New York.
23. Lu, M. *et al.* (1992), "A Parallel Square Root Algorithm for Modified Extended Kalman Filter", *IEEE Transaction on Aerospace and Electronic Systems*, Vol. 28, No. 1, PP. 153-163.
24. Maybeck, P.S. (1979), "*Stochastic Models, Estimation and Control*", Vol.1 Academic Press, Inc. New York.
25. Meinhold, R.J. & Singpurwalla, N.D. (1983), "Understanding the Kalman Filter", *Journal of the American statistician*, Vol.37, No.2, PP.123-126.
26. Melsa, J.L. & Cohn, D.L. (1978), "*Decision and Estimation Theory*", McGraw-Hill, Inc.
27. Morf, M. & Kailath, T. (1975), "Square-Root Algorithms for Least Squares Estimation", *IEEE Trans. Automatic Control*, Vol. AC-20, No. 4, PP. 483-497.

28. Sorenson, H.W. (1970), "Least Square Estimation: From Gauss to Kalman", *IEEE Spectrum*, Vol. 7, PP. 63-68.
29. Strang, G. (1980), "*Linear Algebra and It's Applications*", 2nd Ed., Academic Press, Inc., New York.
30. Vaseghi, S.V. (2000), "*Advanced Digital Signal Processing and Noise Reduction*", John Wiley & Sons, Ltd.
31. Wald, A. (1947), "*Sequential Analysis*", John Wiley & Sons, Inc., New York.
32. Wan, E.A. *et al.* (2000), "Dual Estimation and the Unscented Transformation", *Neural Information Processing Systems*, 12, PP. 666-672, MIT Press.
33. West, M. & Harrison, P.J. (1989), "*Bayesian Forecasting and Dynamic Models*", Springer-Verlag. New York.
34. Wiener, N. (1949), "*The Extrapolation, Interpolation and Smoothing of Stationary Time Series*", John Wiley & Sons, New York.

#### مصادر الإنترنت:

1. Campagnoli, P. *et al.* (2000), "Generalized dynamic linear models for financial time series". URL: <http://www.dma.utc.fr/Pointdoc.htm>.
2. Kybic, J. (1998), "Kalman Filter And Speech Enhancement", Diploma Work. URL: <http://cmp.felk.cvut.cz/~kybic/dipl>.
3. Snyder, R.D. & Forbes, C.S. (2002), "Reconstructing the Kalman Filter for Stationary and Non Stationary Time Series", Unpublished Paper, Dept. of Econometrics & Business Statistics, Monash University, Australia. URL: <http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/2002/wp14-02.pdf>
4. Welch, G. & Bishop, G. (2001), "An Introduction to the Kalman Filter", Course in Kalman Filter, Dept. of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill. URL: <http://www.cs.unc.edu/~tracker/ref/s2001/kalman/index.html>.