Лекция 8.
Задача линейной фильтрации. Фильтр Калмана – Бьюси.
Необходимые сведения о гауссовских векторах.

Рассматриваем пространство R^d, вектор в R^d - это вектор-столбец. Пусть $m \in R^d$ и Q – симметрическая, неотрицательно определённая матрица порядка d (то есть $d \times d$ - матрица). Гауссовский вектор $X \in R^d$ с законом распределения $\mathcal{N}(m, Q)$ - это случайный вектор, преобразование Фурье которого (характеристическая функция) имеет вид

$$E\left(e^{i\lambda^* X}\right) = e^{i\lambda^* m} \cdot e^{-\frac{\lambda^* Q \lambda}{2}}. \qquad (1)$$

Для всех $\lambda \in R^d$. Тогда

$$E(X) = m$$

и

$$Var(X) = Q,$$

где $Var(X)$ обозначает матрицу ковариаций вектора X, которая равна

$$Var(X) = E\left(\left(X - E(X)\right)\left(X - E(X)\right)^*\right).$$

Обратно, если есть X случайный вектор, имеющий м. о. равное m и матрицу ковариаций Q и такой, что все линейные комбинации $\lambda^* X, \lambda \in R^d$, гауссовские, $\lambda^* X \sim \mathcal{N}(\lambda^* m, \lambda^* Q \lambda)$, то, как следует из формулы (1) преобразования Фурье для $\lambda^* X$, вектор $X \sim \mathcal{N}(m, Q)$.
Пусть теперь $X = \binom{U}{V}$, где $U \in R^p, V \in R^q$. Предположим, что для всех $k, j, 1 \le k \le p, 1 \le j \le q$,

$$Cov(U_k, V_j) = E\left(\left(U_k - E(U_k)\right)\left(V_j - E(V_j)\right)^*\right) = 0$$

где U_k, V_j - компоненты U и V. Тогда ковариационная матрица имеет блочный вид

$$Q = \begin{pmatrix} Var(U) & 0 \\ 0 & Var(V) \end{pmatrix}.$$

Следовательно, если $m_U = E(U), m_V = E(V)$, и если $a \in R^p, b \in R^q$, то, поскольку

$$\lambda^* X := \begin{pmatrix} a \\ b \end{pmatrix}^* X = (a^*, b^*) \begin{pmatrix} U \\ V \end{pmatrix} = a^* U + b^* V,$$

имеем

$$Ee^{i(a^* U + b^* V)} = e^{i\lambda^* m} e^{-\frac{\lambda^* Q \lambda}{2}} = e^{i(a^* m_U + b^* m_V)} e^{-\frac{a^* Var(U) a}{2}} e^{-\frac{b^* Var(V) b}{2}} =$$

$$e^{ia^* m_U} \cdot e^{-\frac{a^* Var(U) a}{2}} \cdot e^{ib^* m_V} \cdot e^{-\frac{b^* Var(V) b}{2}} = E\left(e^{ia^* U}\right) \cdot E\left(e^{ib^* V}\right).$$

Таким образом, преобразование Фурье пары (U, V) равно произведению преобразований Фурье U и V, откуда следует независимость U и V. Мы доказали нетривиальную часть (достаточность) следующей теоремы.

Теорема 1. *Если вектор (U, V) гауссовский, то необходимое и достаточное условие независимости U и V состоит в том, что для всех компонент этих векторов $Cov\left(U_k, V_j\right) = 0$, $k = 1, \dots, p$, $j = 1, \dots, q$.*

Рассмотрим теперь гауссовский вектор

$$(X_1, \dots, X_n, Z)$$

и будем искать условный закон распределения Z при условии $\sigma(X_1, \dots, X_n)$. Компоненты $X_1, \dots, X_n, Z$ интегрируемы с квадратом, то есть принадлежат пространству $L^2(\Omega, \mathcal{A}, P)$. Обозначим через H подпространство L^2 конечной размерности, порождённое функциями $X_1, \dots, X_n$ и функцией, равной константе 1. Пусть $\hat{Z}$ - ортогональная проекция Z на H (в пространстве $L^2(\Omega, \mathcal{A}, P)$, снабжённом скалярным произведением $\langle U, V \rangle = E(UV)$).

Лемма 1. *Случайная величина $Z - \hat{Z}$ не зависит от $(X_1, \dots, X_n)$, имеет гауссовское распределение со средним 0 и ортогональная проекция Z* на *H почти наверное совпадает с условным математическим ожиданием*

$$\hat{Z} = E(Z | \sigma(X_1, \dots, X_n)).$$

Доказательство. По определению проекции в L^2, вектор $\hat{Z}$ - это единственный вектор вида $\hat{Z} = 1 \cdot \lambda_0 + \sum_{k=1}^{n} \lambda_k \cdot X_k$ такой, что

$$\langle Z-\hat{Z},1\rangle = E\big[\big(Z-\hat{Z}\big)\cdot 1\big] = E\big(Z-\hat{Z}\big) = 0.$$

и для $k = 1,2,\dots n$

$$\langle Z-\hat{Z},X_k\rangle = E\left(\big(Z-\hat{Z}\big)\cdot X_k\right) = 0.$$

Таким образом, $\forall k = 1,2,\dots,n$

$$Cov\big(Z-\hat{Z},X_k\big) = E\left(\big(Z-\hat{Z}\big)\cdot X_k\right) - E\big(Z-\hat{Z}\big)\cdot EX_k = 0.$$

Независимость $Z-\hat{Z}$ от $(X_1,\dots,X_n)$ следует теперь из Теоремы 1. Наконец,

$$E\big(Z|\sigma(X_1,\dots,X_n)\big) = E\left(\big(Z-\hat{Z}\big)|\sigma(X_1,\dots,X_n)\right) +$$

$$E\left(\hat{Z}|\sigma(X_1,\dots,X_n)\right) = \hat{Z}$$

Первое слагаемое средней части последнего равенства равно нулю, так как $Z-\hat{Z}$ не зависит от $\sigma(X_1,\dots,X_n)$ и $E\big(Z-\hat{Z}\big) = 0$. Во втором слагаемом, поскольку $\hat{Z}$ - ортогональная проекция Z на H, $\hat{Z}$ есть линейная комбинация $(1,X_1,\dots,X_n)$, поэтому $\hat{Z}$ измерима относительно $\sigma(X_1,\dots,X_n)$ и $E\left(\hat{Z}\middle|\sigma(X_1,\dots,X_n)\right) = \hat{Z}$. Кроме того

$$Var\big(Z-\hat{Z}\big) = E\big(Z-\hat{Z}\big)^2.$$

Предложение 1. *Условный закон распределения* Z *при условии* $\sigma(X_1,\dots,X_n)$ *является гауссовским с (условным) средним* $\hat{Z}$, *где* $\hat{Z}$ – *ортогональная проекция* Z *на векторное пространство, порождённое* $1,X_1,\dots,X_n$.

Доказательство. Представим $Z = \hat{Z} + Z - \hat{Z}$. Тогда из независимости $Z-\hat{Z}$ от $\sigma(X_1,\dots,X_n)$ следует

$$E\left(e^{i\lambda Z}|\sigma(X_1,\dots,X_n)\right) = e^{i\lambda\hat{Z}}E\left(e^{i\lambda(Z-\hat{Z})}|\sigma(X_1,\dots,X_n)\right) =$$

$$e^{i\lambda\hat{Z}}E\big(e^{i\lambda(Z-\hat{Z})}\big).$$

Случайная величина $Z - \hat{Z}$ - гауссовская, как линейная комбинация компонент гауссовского вектора $(X_1, \dots, X_n, Z)$. Таким образом, в правой части последнего равенства стоит характеристическая функция с.в., имеющей нормальное распределение со средним $\hat{Z}$.

Фильтр Калмана – Бьюси.

Постановка задачи. Для $t \geq 1$ рассмотрим уравнения

$$X_{t+1} = AX_t + \varepsilon_{t+1}$$

$$Y_t = CX_t + \tau_t \qquad (2)$$

где X_t и ε_t случайные векторы в R^d, Y_t и τ_t случайные векторы в R^p, $A - d \times d$ матрица и C - $p \times d$ матрица. Предположим, что:

- $X_1, \varepsilon_1, \varepsilon_2, \dots \tau_1, \tau_2, \dots$ независимые **гауссовские** векторы
- ε_t центрированы, с матрицей ковариаций Q
- τ_t центрированы, с **обратимой** матрицей ковариаций R
- Матрицы A и C **детерминированы**.

Первое из уравнений (2) называется ***уравнением состояний***, а второе – ***уравнением наблюдений***. Наблюдаемыми являются вектора Y_t, которые поступают последовательно в моменты времени $t = 1,2, \dots$Вектора X_t ***ненаблюдаемые.*** По наблюдениям Y_t требуется построить оценки для векторов X_t и оценить точность оценок. При этом оценки динамически пересчитываются с поступлением каждого нового наблюдения Y_t. Естественной оценкой ненаблюдаемого гауссовского вектора X_t по доступной наблюдаемой информации $Y_1, Y_2, \dots, Y_t$ является, как мы видели, условное математическое ожидание $E(X_t | \sigma(Y_1, Y_2, \dots, Y_t))$ и с вычислительной точки зрения целью фильтра Калмана – Бьюси является получение рекуррентных формул для вычисления проекции

$$\hat{X}_t = E(X_t | \sigma(Y_1, Y_2, \dots, Y_t))$$

и точности прогноза

$$P_t := Var(X_t - \hat{X}_t),$$

где через Var обозначена ковариационная матрица вектора ошибки прогноза $X_t - \hat{X}_t$. Напомним, что условное м.о. $E(X_t|\sigma(Y_1, \dots, Y_t))$ ***вектора*** относительно σ – алгебры – это ***случайный вектор***, компоненты которого равны условным м.о. соответствующих компонент вектора X_t. Рекуррентные формулы должны позволять пересчитывать $\hat{X}_t$ и P_t по мере поступления новой информации.

Разложение с помощью обновлений.

Введём:

- Наилучшую аппроксимацию $\hat{X}_t^-$, которую можно получить для X_t ***к моменту*** $\boldsymbol{t-1}$, то есть ***до*** поступления наблюдения в момент t:

$$\hat{X}_t^- := E(X_t|\sigma(Y_1, \dots, Y_{t-1}))$$

- Матрицу ковариаций P_t^- возникающей при этом ошибки

$$P_t^- := Var(X_t - \hat{X}_t^-)$$

- Обновление (информации) J_t, привносимое новым поступившим наблюдением Y_t

$$J_t := Y_t - E\big(Y_t\big|\sigma(Y_1, \dots, Y_{t-1})\big).$$

(заметим, что J_t не зависит от $Y_1, Y_2, \dots, Y_{t-1}$ (по Лемме 1) и, следовательно, действительно является той частью Y_t, которая приносит новую информацию, которая ***«обновляет»***). Следующая лемма легко доказывается.

Лемма 2. *Верны следующие соотношения*

$$\hat{X}_{t+1}^- = A\hat{X}_t,$$

$$J_t = Y_t - C\hat{X}_t^- = C\big(X_t - \hat{X}_t^-\big) + \tau_t,$$

$$P_{t+1}^- = AP_tA^* + Q.$$

Доказательство. Имеем

$$\hat{X}_{t+1}^{-} = E\big(X_{t+1}\big|\sigma(Y_1, \dots, Y_t)\big) = E\big((AX_t + \varepsilon_{t+1})\big|\sigma(Y_1, \dots, Y_t)\big) = A\hat{X}_t.$$

Мы воспользовались тем, что

$$\sigma(Y_1, \dots, Y_t) \subset \sigma(X_1, \varepsilon_1, \dots, \varepsilon_t, \tau_1, \tau_2, \dots, \tau_t)$$

и предположениями модели, из которых следует, что ε_{t+1} не зависит от $\sigma(Y_1, \dots, Y_t)$ и, следовательно,

$$E\big(\varepsilon_{t+1}\big|\sigma(Y_1, \dots, Y_t)\big) = E\varepsilon_{t+1} = 0.$$

Второе равенство леммы очевидным образом следует из (2) и предположений модели. Для доказательства третьего равенства заметим, что

$$P_{t+1}^{-} = Var\left(X_{t+1} - E\big(X_{t+1}\big|\sigma(Y_1, \dots, Y_t)\big)\right) = Var[AX_t + \varepsilon_{t+1} -$$

$$E((AX_t + \varepsilon_{t+1})|\sigma(Y_1, \dots, Y_t))] = Var(A\big(X_t - \hat{X}_t\big) + \varepsilon_{t+1}) =$$

$$AP_tA^* + Q. \tag{3}$$

(в последнем равенстве воспользовались независимостью векторов $A(X_t - \hat{X}_t)$ и ε_{t+1}).

Нам понадобится также следующая лемма из теории гильбертовых пространств.

Лемма 3. *Пусть H_1 и H_2 два замкнутых векторных подпространства гильбертова пространства $\mathfrak{H}$, которые ортогональны между собой. Пусть $H = H_1 \oplus H_2$. Обозначим π, π_1, π_2 ортогональные проекции на H, H_1, H_2. Тогда*

$$\pi = \pi_1 + \pi_2.$$

Заметим теперь, что каждая компонента вектора J_t

$$J_t = Y_t - C\hat{X}_t^{-} = C\big(X_t - \hat{X}_t^{-}\big) + \tau_t$$

ортогональна любой из компонент векторов $1, Y_1, \dots, Y_{t-1}$, напомним, что $Y_t = CX_t + \tau_t$, то есть (по Теореме 1) J_t не зависит от $\sigma(Y_1, \dots, Y_{t-1})$. Кроме того, из определения J_t следует, что

$$\sigma(Y_1, \dots, Y_{t-1}, Y_t) = \sigma(Y_1, \dots, Y_{t-1}, J_t).$$

Из Леммы 3, где в качестве H_1 берется L^2, порождённое векторами $1, Y_1, \dots, Y_{t-1}$, а в качестве H_2 - подпространство в L^2, порождённое вектором J_t , следует, что ортогональная проекция X_t на подпространство в L^2, порождённое всеми компонентами векторов $1, Y_1, \dots, Y_t$ (то есть $\hat{X}_t$) равна ортогональной проекции X_t на подпространство в L^2, порождённое всеми компонентами векторов $1, Y_1, \dots, Y_{t-1}$ (то есть $\hat{X}_t^-$) плюс ортогональная проекция X_t на «одномерное» подпространство в L^2, порождённое вектором J_t (то есть $E(X_t | J_t = (J_{1t}, J_{2t}, \dots, J_{pt})^*)$). Это последнее условное математическое ожидание есть, по определению, d - мерный вектор, компоненты которого равны условным математическим ожиданиям компонент вектора X_t относительно $\sigma\big(J_{1t}, J_{2t}, \dots, J_{pt}\big)$. Так как все рассматриваемые вектора гауссовские, то все компоненты вектора $E(X_t | J_t = \big(J_{1t}, J_{2t}, \dots, J_{pt}\big)^*)$ есть линейные комбинации компонент $\big(J_{1t}, J_{2t}, \dots, J_{pt}\big)$, то есть

$$E\left(X_t \middle| \sigma\big(J_{1t}, \dots, J_{pt}\big)\right) = K_t J_t,$$

где K_t - матрица порядка $d \times p$, называемая ***матрицей коэффициентов усиления Калмана***. Мы получили важное соотношение

$$\hat{X}_t = \hat{X}_t^- + K_t J_t, \qquad (4)$$

где $\hat{X}_t^-$ и $K_t J_t$ ортогональны и независимы.

Вычисление матрицы коэффициентов усиления Калмана.

По определению ортогональной проекции, компоненты случайных векторов

$$X_t - E(X_t | \sigma\big(J_{1t}, \dots, J_{pt}\big)) = X_t - K_t J_t$$

и J_t ортогональны. Отсюда

$$E\big((X_t - K_t J_t) \cdot J_t^*\big) = 0.$$

То есть

$$E(X_t J_t^*) = K_t E(J_t J_t^*). \qquad (5)$$

Поскольку J_t центрирован и не зависит от $\sigma(Y_1, \dots, Y_{t-1})$, получим, используя второе соотношение Леммы 2, (4), (5)

$$E(X_t J_t^*) = E\left(\left(X_t - \hat{X}_t^-\right) J_t^*\right) + E\left(\hat{X}_t^- J_t^*\right) = E\left(\left(X_t - \hat{X}_t^-\right) J_t^*\right) =$$

$$E\left(\left(X_t - \hat{X}_t^-\right)\left(C\left(X_t - \hat{X}_t^-\right) + \tau_t\right)^*\right) = P_t^- C^*.$$

и

$$E(J_t J_t^*) = E\left(\left(C\left(X_t - \hat{X}_t^-\right) + \tau_t\right)\left(C\left(X_t - \hat{X}_t^-\right) + \tau_t\right)^*\right) =$$

$$CE\left(\left(X_t - \hat{X}_t^-\right)\left(X_t - \hat{X}_t^-\right)^*\right) C^* + E(\tau_t \tau_t^*) = CP_t^- C^* + R.$$

Отсюда и из (5) найдём K_t (напомним, что R обратима)

$$K_t = E(X_t J_t^*)[E(J_t J_t^*)]^{-1} = P_t^- C^* (CP_t^- C^* + R)^{-1}. \qquad (6)$$

Наконец, покажем как переходить к P_t от P_t^-, используя Лемму 2 и соотношение $\hat{X}_t = \hat{X}_t^- + K_t J_t$

$$P_t = E\left((X_t - \hat{X}_t)(X_t - \hat{X}_t)^*\right) =$$

$$E\left((X_t - \hat{X}_t^- - K_t J_t)(X_t - \hat{X}_t^- - K_t J_t)^*\right) =$$

$$E\left((X_t - \hat{X}_t^-)(X_t - \hat{X}_t^-)^*\right) - E\left(K_t J_t (X_t - \hat{X}_t^-)^*\right) -$$

$$E\left((X_t - \hat{X}_t^-)(K_t J_t)^*\right) + E[(K_t J_t)(K_t J_t)^*] =$$

$$P_t^- - K_t E\left(\left(C\left(X_t - \hat{X}_t^-\right) + \tau_t\right)(X_t - \hat{X}_t^-)^*\right)$$

$$-E(X_t (K_t J_t)^*) + E\left(\hat{X}_t^- (K_t J_t)^*\right)$$

$$+E((K_t J_t)(K_t J_t)^*) =$$

$$P_t^- - K_t CE\left((X_t - \hat{X}_t^-)(X_t - \hat{X}_t^-)^*\right) - E(X_t (K_t J_t)^*) +$$

$$E\left(\hat{X}_t^-(K_tJ_t)^*\right) + K_tE(J_tJ_t^*)K_t^* = P_t^- - K_tCP_t^-, \tag{7}$$

где в последнем равенстве мы воспользовались ортогональностью $\hat{X}_t^-$ и K_tJ_t (см. (4)) и равенством $K_tE(J_tJ_t^*) = E(X_tJ_t^*)$ (см. (5)).

Начало рекурсии.

Рекурсия начинается с вектора $\hat{X}_1^-$, который есть просто проекция X_1 на одномерное пространство констант (так как наблюдений Y_t к этому моменту ещё нет), то есть $\hat{X}_1^- = E(X_1)$. Соответственно,

$$P_1^- = Var\left(X_1 - \hat{X}_1^-\right) = Var\left(X_1 - E(X_1)\right).$$

Существует другой, достаточно общий метод выбора начальных условий. Если все собственные числа матрицы A в уравнении состояний лежат внутри единичного круга, то VAR (1) процесс в системе Калмана является, как мы знаем, стационарным в широком смысле, т.е. его средние и ковариации постоянны, не зависят от времени. Переходя к безусловным математическим ожиданиям, получим $E(X_{t+1}) = AE(X_t)$, откуда, в силу стационарности,

$$E(X_{t+1}) = E(X_t) \Leftrightarrow (I_d - A)E(X_t) = 0.$$

По предположению, единица не является собственным числом матрицы A, то есть матрица $(I_d - A)$ обратима и уравнение имеет единственное решение $E(X_t) = 0$. Ковариационная матрица Σ_t (безусловная) вектора X_t также легко вычисляется. Из уравнения состояний получим

$$\Sigma_{t+1} = A\Sigma_tA^* + Q.$$

Обозначая общую для всех t ковариационную матрицу через Σ, получим уравнение

$$\Sigma = A\Sigma A^* + Q.$$

Решение этого уравнение даётся формулой (см. Добавление в конце лекции)

$$vec(\Sigma) = [I_{d^2} - (A \otimes A)]^{-1} \cdot vec(Q).$$

Таким образом, если все собственные числа матрицы A лежат внутри единичного круга, то итерации фильтра Калмана могут начинаться с $\hat{X}_1^- = E(X_1) = 0$ и с $d \times d$ матрицы P_1^- такой, что

$$vec(P_1^-) = [I_{d^2} - (A \otimes A)]^{-1} \cdot vec(Q).$$

Если же собственные числа матрицы A могут лежать на единичном круге или вне единичного круга, или если X_1 не может рассматриваться как значение случайного процесса, удовлетворяющего уравнению состояний, то выбор начала рекурсии диктуется спецификой задачи и опытом того, кто применяет метод.

Алгоритм.

Подводя итог нашим вычислениям, приходим к следующему алгоритму. Имея начальные значения $\hat{X}_1^-$ и P_1^-, мы имеем поступившее первое наблюдение Y_1 и хотим вычислить $\hat{X}_2^-$ и P_2^-, затем поступает второе наблюдение Y_2 мы хотим найти $\hat{X}_3^-$ и P_3^- и так далее до $\hat{X}_T^-$ и P_T^-. Покажем в общем виде, как зная $\hat{X}_t^-$ и P_t^- и получив следующее наблюдение Y_t, найти $\hat{X}_{t+1}^-$ и P_{t+1}^-. Из (3), (7) и (6) получим переход от P_t^- к P_{t+1}^- :

$$P_{t+1}^- = AP_tA^* + Q = A\underbrace{\left(P_t^- - K_tCP_t^-\right)}_{\square_t}A^* + Q =$$

$$A\left(P_t^- - \underbrace{P_t^-C^*(CP_t^-C^* + R)^{-1}}_{K_t}CP_t^-\right)A^* + Q. \qquad (8)$$

Из (4) и Леммы 2 получим переход от $\hat{X}_t^-$ к $\hat{X}_{t+1}^-$:

$$\hat{X}_t = \hat{X}_t^- + K_tJ_t = \hat{X}_t^- + K_t\underbrace{\left(Y_t - C\hat{X}_t^-\right)}_{J_t}. \qquad (9)$$

Умножая обе части последнего равенства на A и снова пользуясь Леммой 2, получим

$$\hat{X}_{t+1}^{-} = A\hat{X}_t = A\hat{X}_t^{-} + AK_t\left(Y_t - C\hat{X}_t^{-}\right) =$$

$$A\hat{X}_t^{-} + A\,\underbrace{P_t^{-}C^{*}(CP_t^{-}C^{*} + R)^{-1}}_{K_t}\left(Y_t - C\hat{X}_t^{-}\right) \qquad (10)$$

Таким образом, формулы (8) и (10) дают искомый переход от P_t^{-} и $\hat{X}_t^{-}$ к P_{t+1}^{-} и $\hat{X}_{t+1}^{-}$. На каждом шаге алгоритма из (4) и (7) также находим проекции $\hat{X}_t$ и дисперсии ошибок P_t.

Уравнение Риккати. Напомним уравнение Риккати из теории управления линейными системами (Теорема 3 Лекции 8):

$$\Gamma_t = \rho_{(A,B)}(\Gamma_{t+1}),$$

где

$$\rho_{(A,B)}(\Gamma) = Q + A^{*}\Gamma A - (A^{*}\Gamma B)(R + B^{*}\Gamma B)^{-1}(B^{*}\Gamma A).$$

Рассмотрим уравнение, которому удовлетворяет P_t^{-}

$$P_{t+1}^{-} = AP_tA^{*} + Q = A\,\underbrace{(P_t^{-} - K_tCP_t^{-})}_{P_t}\,A^{*} + Q = Q +$$

$$AP_t^{-}A^{*} - AK_tCP_t^{-}A^{*},$$

где

$$K_t = P_t^{-}C^{*}(CP_t^{-}C^{*} + R)^{-1}.$$

Поэтому P_{t+1}^{-} можно переписать также в виде

$$P_{t+1}^{-} = Q + AP_t^{-}A^{*} - AP_t^{-}C^{*}(CP_t^{-}C^{*} + R)^{-1}CP_t^{-}A^{*} = \rho_{(A^{*},C^{*})}(P_t^{-}).$$

То есть P_t^{-} удовлетворяет уравнению Риккати с (A^{*}, C^{*}).

Добавление. Напомним некоторые факты и обозначения из теории матриц. Пусть A - $m \times n$ матрица и B - $p \times q$ матрица.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1q} \\ \vdots & \ddots & \vdots \\ b_{p1} & \dots & b_{pq} \end{pmatrix}$$

Тогда $mp \times nq$ матрица их тензорного (или кронекеровского) произведения обозначается $A \otimes B$ и равна

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}$$

Далее, если A - $m \times n$ матрица, то $vec(A)$ – это вектор-столбец размерности mn, получающийся составлением всех столбцов матрицы (слева направо) в один столбец, каждый следующий столбец располагается ниже предыдущего. Имеет место следующее полезное соотношение, доказательство которого можно найти в книге Hamilton (Appendix 10.A)

Предложение. *Пусть A,B и C – матрицы, размерности которых таковы, что произведение ABC существует. Тогда*

$$vec(ABC) = (C^* \otimes A) \cdot vec(B).$$

Применяя это Предложение к уравнению $\Sigma = A\Sigma A^* + Q$,

получим, применяя оператор vec к обеим частям этого уравнения

$$vec(\Sigma) = vec(A\Sigma A^*) + vec(Q) = (A \otimes A) \cdot vec(\Sigma) + vec(Q),$$

откуда

$$\left(I_{d^2} - (A \otimes A)\right) \cdot vec(\Sigma) = vec(Q),$$

$$vec(\Sigma) = \left(I_{d^2} - (A \otimes A)\right)^{-1} \cdot vec(Q).$$

Собственные числа матрицы $A \otimes A$ равны $\lambda_i \lambda_j, i, j = 1,2, \dots, d,$ где $\lambda_i, i = 1, \dots, d,$ собственные числа матрицы A. Поэтому они все тоже лежат внутри единичного круга, и матрица $\left(I_{d^2} - (A \otimes A)\right)$ обратима.