Приложение Г

Оптимизация процедур согласования индустриальных данных за счет аналитического решения соответствующей задачи

Как отмечалось в основном отчете для оптимизации по скорости выполнения вычислений есть возможность использовать аналитическую модель, построенную на теореме Ефремова-Козлова, изложенной в [1, 2]. Получаемые с ее помощью формулы для согласования результатов неточных индустриальных данных допускают прямой и непосредственный матричный расчет без необходимости решения оптимизационной задачи, поскольку основаны на условиях Каруша-Куна-Таккера задачи условной оптимизации. Полученные соотношения оптимизируют расчеты по времени (в смысле количества обращений к математической модели пользователя, сокращая их до одного – но, для некоторых из рассмотренных постановок, с условием возвращения не только значений математической модели, но и ее частных производных). Задача учета неопределенности исходных данных также оптимизируется, поскольку при выполненном аналитическом решении оказывается возможным получить оценки точности конечных результатов согласования, основываясь только на тех же самых значениях, возвращенных математической моделью (и значениях частных производных).

Переформулирование [1,2] для задачи согласования индустриальных данных

Пусть $\mathbf{x}$ – вектор результатов измерений, $\mathbf{\Delta x}$ – вектор абсолютных погрешностей, $\mathbf{x}^*$ – значения измеряемых величин (с которыми в идеале должны совпасть результаты согласования). Результат Ефремова-Козлова показывает, что

$$\mathbf{x}^* = \arg \min_{\substack{\mathbf{x}: A\cdot\mathbf{x}=\mathbf{b}, \\ (\mathbf{x}-\mathbf{c})^T\cdot(\mathbf{x}-\mathbf{c})\le r^2}} \|\mathbf{x}\|_2^2$$

$$\Downarrow \mathbf{x}^* = \mathbf{x} - \mathbf{\Delta x}$$

$$\mathbf{x}^* = \mathbf{x} - \arg \min_{\substack{\mathbf{\Delta x}: A\cdot\mathbf{\Delta x}=A\cdot\mathbf{x}-\mathbf{b}, \\ (\mathbf{\Delta x}-(\mathbf{x}-\mathbf{c}))^T\cdot(\mathbf{\Delta x}-(\mathbf{x}-\mathbf{c}))\le r^2}} \|\mathbf{\Delta x}\|_2^2,$$

что эквивалентно выполнению условий вида $A \cdot \mathbf{x}^* = \mathbf{b}$ и $(\mathbf{x}^* - \mathbf{c})^T \cdot (\mathbf{x}^* - \mathbf{c}) \le r^2$ для результатов согласования (точных значений измеряемых величин). Здесь $\|\mathbf{x}\|_2^2 = \mathbf{x}^T \cdot \mathbf{x}$ – Евклидова норма вектора $\boldsymbol{x}$.

Следуя [1], аналитическое решение имеет вид

$$\mathbf{x}^* = \mathbf{x} + P \cdot (A \cdot \mathbf{x} - \mathbf{b}) + \lambda \cdot (1 + \lambda)^{-1} \cdot D \cdot (\mathbf{x} - \mathbf{c}),$$

где n – размер векторов $\mathbf{x}$ и $\mathbf{c}$; $D = E_n - A^T \cdot (A \cdot A^T)^{-1} \cdot A$; E_n – единичная матрица размером $n \times n$; $P = A^T \cdot (A \cdot A^T)^{-1}$; множители Лагранжа λ определяются решением квадратного уравнения

$$a \cdot \lambda^2 + h \cdot \lambda + d = 0,$$

$$a = (A \cdot \mathbf{x} - \mathbf{b})^T \cdot P^T \cdot P \cdot (A \cdot \mathbf{x} - \mathbf{b}) - (A \cdot \mathbf{x} - \mathbf{b})^T \cdot P^T \cdot (\mathbf{x} - \mathbf{c}) - (\mathbf{x} - \mathbf{c})^T \cdot P \cdot (A \cdot \mathbf{x} - \mathbf{b}) + (\mathbf{x} - \mathbf{c})^T \cdot (\mathbf{x} - \mathbf{c}) - (\mathbf{x} - \mathbf{c})^T \cdot D \cdot (\mathbf{x} - \mathbf{c}) - r^2,$$

$$h = 2 \cdot a,\ d = a + (\mathbf{x} - \mathbf{c})^T \cdot D \cdot (\mathbf{x} - \mathbf{c}).$$

Учет многих неравенств осуществляется заменой условия $(\mathbf{x}-\mathbf{c})^T \cdot (\mathbf{x}-\mathbf{c}) \le r^2$ на условие $Q \cdot \mathbf{x}^* \le \mathbf{t}$, где *Q* – матрица *mxn*, *m* – размер вектора **t**,

Выпишем вслед за [2] функцию Лагранжа для данной задачи:

$$L = \|\mathbf{x}\|_2^2 + \boldsymbol{\lambda}_\mathbf{0} \cdot (A \cdot \mathbf{x} - \mathbf{b}) + \boldsymbol{\lambda} \cdot (Q \cdot \mathbf{x} - \mathbf{t}).$$

Необходимыми условиями достижения оптимума в соответствии с условиями Каруша-Куна-Таккера являются равенства нулю всех соответствующих частных производных:

$$\frac{\partial L}{\partial \mathbf{x}} = 2 \cdot \mathbf{x} + A^T \cdot \boldsymbol{\lambda}_\mathbf{0} + Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_n,$$

$$\frac{\partial L}{\partial \boldsymbol{\lambda}_\mathbf{0}} = A \cdot \mathbf{x} - \mathbf{b} = \mathbf{0}_m,$$

$$\frac{\partial L}{\partial \boldsymbol{\lambda}} = Q \cdot \mathbf{x} - \mathbf{t} = \mathbf{0}_k.$$

Решение ищем следующим образом.

Подстановка второго уравнения в первое уравнение:

$2 \cdot \mathbf{x} + A^T \cdot \boldsymbol{\lambda}_\mathbf{0} + Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_n$

$2 \cdot A \cdot \mathbf{x} + A \cdot A^T \cdot \boldsymbol{\lambda}_\mathbf{0} + A \cdot Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_m$

$2 \cdot \mathbf{b} + A \cdot A^T \cdot \boldsymbol{\lambda}_\mathbf{0} + A \cdot Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_m$

$A \cdot Q^T \cdot \boldsymbol{\lambda} = -2 \cdot \mathbf{b} - A \cdot A^T \cdot \boldsymbol{\lambda}_\mathbf{0}$

$(A \cdot Q^T)^T \cdot A \cdot Q^T \cdot \boldsymbol{\lambda} = -2 \cdot (A \cdot Q^T)^T \cdot \mathbf{b} - (A \cdot Q^T)^T \cdot A \cdot A^T \cdot \boldsymbol{\lambda}_\mathbf{0}$

$Q \cdot A^T \cdot A \cdot Q^T \cdot \boldsymbol{\lambda} = -2 \cdot Q \cdot A^T \cdot \mathbf{b} - Q \cdot A^T \cdot A \cdot A^T \cdot \boldsymbol{\lambda}_\mathbf{0}$

$\boldsymbol{\lambda} = -2 \cdot (Q \cdot A^T \cdot A \cdot Q^T)^{-1} \cdot Q \cdot A^T \cdot \mathbf{b} - (Q \cdot A^T \cdot A \cdot Q^T)^{-1} \cdot Q \cdot A^T \cdot A \cdot A^T \cdot \boldsymbol{\lambda}_\mathbf{0} \rightarrow$ kx1

или же выразим $\boldsymbol{\lambda}_\mathbf{0}$:

$-A \cdot Q^T \cdot \boldsymbol{\lambda} - 2 \cdot \mathbf{b} = A \cdot A^T \cdot \boldsymbol{\lambda}_\mathbf{0}$, матрица $A \cdot A^T$ является квадратной,

$$\boldsymbol{\lambda}_\mathbf{0} = -(A \cdot A^T)^{-1} \cdot A \cdot Q^T \cdot \boldsymbol{\lambda} - 2 \cdot (A \cdot A^T)^{-1} \cdot \mathbf{b}$$

Подстановка третьего уравнения в первое уравнение:

$2 \cdot \mathbf{x} + A^T \cdot \boldsymbol{\lambda}_\mathbf{0} + Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_n$

$2 \cdot Q \cdot \mathbf{x} + Q \cdot A^T \cdot \boldsymbol{\lambda}_\mathbf{0} + Q \cdot Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_k$

$2 \cdot \mathbf{t} + Q \cdot A^T \cdot \boldsymbol{\lambda}_\mathbf{0} + Q \cdot Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_k$

Выразим вектор $\boldsymbol{\lambda}_\mathbf{0}$ и приравняем к полученному выше уравнению – так и найдем вектор $\boldsymbol{\lambda}$, который подставим потом в первое уравнение

$(Q \cdot A^T)^T \cdot Q \cdot A^T \cdot \boldsymbol{\lambda}_\mathbf{0} = -2 \cdot (Q \cdot A^T)^T \cdot \mathbf{t} - (Q \cdot A^T)^T \cdot Q \cdot Q^T \cdot \boldsymbol{\lambda}$

$A \cdot Q^T \cdot Q \cdot A^T \cdot \boldsymbol{\lambda}_\mathbf{0} = -2 \cdot A \cdot Q^T \cdot \mathbf{t} - A \cdot Q^T \cdot Q \cdot Q^T \cdot \boldsymbol{\lambda}$

$$\boldsymbol{\lambda}_\mathbf{0} = -2 \cdot (A \cdot Q^T \cdot Q \cdot A^T)^{-1} \cdot A \cdot Q^T \cdot \mathbf{t} - (A \cdot Q^T \cdot Q \cdot A^T)^{-1} \cdot A \cdot Q^T \cdot Q \cdot Q^T \cdot \boldsymbol{\lambda}$$

Приравняем получившиеся выражения для $\boldsymbol{\lambda}_\mathbf{0}$:

$$-2 \cdot (A \cdot Q^T \cdot Q \cdot A^T)^{-1} \cdot A \cdot Q^T \cdot \mathbf{t} - (A \cdot Q^T \cdot Q \cdot A^T)^{-1} \cdot A \cdot Q^T \cdot Q \cdot Q^T \cdot \boldsymbol{\lambda} \\ = -(A \cdot A^T)^{-1} \cdot A \cdot Q^T \cdot \boldsymbol{\lambda} - 2 \cdot (A \cdot A^T)^{-1} \cdot \mathbf{b}$$

$$((A\cdot A^T)^{-1}\cdot A\cdot Q^T-(A\cdot Q^T\cdot Q\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot Q\cdot Q^T)\cdot\boldsymbol{\lambda}$$
$$=2\cdot(A\cdot Q^T\cdot Q\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot\mathbf{t}-2\cdot(A\cdot A^T)^{-1}\cdot\mathbf{b}$$

$$((A\cdot A^T)^{-1}\cdot A\cdot Q^T-(A\cdot Q^T\cdot Q\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot Q\cdot Q^T)\cdot\boldsymbol{\lambda}$$
$$=2\cdot(A\cdot Q^T\cdot Q\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot\mathbf{t}-2\cdot(A\cdot A^T)^{-1}\cdot\mathbf{b}$$

Пусть $H=(A\cdot A^T)^{-1}\cdot A\cdot Q^T-(A\cdot Q^T\cdot Q\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot Q\cdot Q^T$.

Тогда
$$H^T\cdot H\cdot\boldsymbol{\lambda}=2\cdot H^T\cdot(A\cdot Q^T\cdot Q\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot\mathbf{t}-2\cdot H^T\cdot(A\cdot A^T)^{-1}\cdot\mathbf{b}$$

$$\boldsymbol{\lambda}=2\cdot(H^T\cdot H)^{-1}\cdot H^T\cdot(A\cdot Q^T\cdot Q\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot\mathbf{t}-2\cdot(H^T\cdot H)^{-1}\cdot H^T\cdot(A\cdot A^T)^{-1}\cdot\mathbf{b}$$

Подставим получившееся выражение для $\boldsymbol{\lambda}$ в одно из выражений для $\boldsymbol{\lambda_0}$ (все равно какое – ответ получается одинаковый):

$$\boldsymbol{\lambda_0}=-(A\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot\boldsymbol{\lambda}-2\cdot(A\cdot A^T)^{-1}\cdot\mathbf{b}$$
$$=-2\cdot(A\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot(H^T\cdot H)^{-1}\cdot H^T\cdot(A\cdot Q^T\cdot Q\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot\mathbf{t}+$$

$$2\cdot(A\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot H^T\cdot(A\cdot A^T)^{-1}\cdot\mathbf{b}-2\cdot(A\cdot A^T)^{-1}\cdot\mathbf{b}=$$

$$-2\cdot(A\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot(H^T\cdot H)^{-1}\cdot H^T\cdot(A\cdot Q^T\cdot Q\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot\mathbf{t}+$$

$$2\cdot(A\cdot A^T)^{-1}\cdot(A\cdot Q^T\cdot H^T\cdot(A\cdot A^T)^{-1}-E_m)\cdot\mathbf{b}$$

Подставим всё в первое уравнение, получим

$$2\cdot\mathbf{x}+A^T\cdot\boldsymbol{\lambda_0}+Q^T\cdot\boldsymbol{\lambda}=\mathbf{0}_n$$

$$\mathbf{x}=-\frac{1}{2}\cdot A^T\cdot\boldsymbol{\lambda_0}-\frac{1}{2}\cdot Q^T\cdot\boldsymbol{\lambda}$$

$$\mathbf{x}=A^T\cdot(A\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot(H^T\cdot H)^{-1}\cdot H^T\cdot(A\cdot Q^T\cdot Q\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot\mathbf{t}-A^T\cdot(A\cdot A^T)^{-1}$$
$$\cdot(A\cdot Q^T\cdot H^T\cdot(A\cdot A^T)^{-1}-E_m)\cdot\mathbf{b}-Q^T\cdot(H^T\cdot H)^{-1}\cdot H^T\cdot(A\cdot Q^T\cdot Q\cdot A^T)^{-1}$$
$$\cdot A\cdot Q^T\cdot\mathbf{t}+Q^T\cdot(H^T\cdot H)^{-1}\cdot H^T\cdot(A\cdot A^T)^{-1}\cdot\mathbf{b}$$

вынесем общие множители и получим ответ:

$$\mathbf{x}=W_{\mathbf{t}}\cdot\mathbf{t}+W_{\mathbf{b}}\cdot\mathbf{b}$$

где $W_{\mathbf{t}}=A^T\cdot(A\cdot A^T)^{-1}\cdot A\cdot Q^T\cdot(H^T\cdot H)^{-1}\cdot H^T\cdot(A\cdot Q^T\cdot Q\cdot A^T)^{-1}\cdot A\cdot Q^T-Q^T\cdot(H^T\cdot H)^{-1}\cdot H^T\cdot(A\cdot Q^T\cdot Q\cdot A^T)^{-1}\cdot A\cdot Q^T=$

$(A^T\cdot(A\cdot A^T)^{-1}\cdot A-E_n)\cdot Q^T\cdot(H^T\cdot H)^{-1}\cdot H^T\cdot(A\cdot Q^T\cdot Q\cdot A^T)^{-1}\cdot A\cdot Q^T$,

$W_{\mathbf{b}}=Q^T\cdot(H^T\cdot H)^{-1}\cdot H^T\cdot(A\cdot A^T)^{-1}-A^T\cdot(A\cdot A^T)^{-1}\cdot(A\cdot Q^T\cdot H^T\cdot(A\cdot A^T)^{-1}-E_m)$.

Учет нелинейных ограничений в виде уравнений.

В линейной постановке мы имели ограничения вида

$$A\cdot\mathbf{x}^*=\mathbf{b}$$

при $\mathbf{x}^*=\mathbf{x}-\boldsymbol{\Delta}\mathbf{x}$.

Допустим, что имеем дело с нелинейным ограничением вида

$$\mathbf{g}(\mathbf{x}^*) = \mathbf{0}_m,$$

где **g** – нелинейная вектор-функция.

Полагая погрешности и, соответственно, разности между результатами измерениями и результатами согласования малыми, т.е. $||\mathbf{\Delta x}|| \ll 1$, получим:

$$\mathbf{g}(\mathbf{x} - \mathbf{\Delta x}) = \mathbf{g}(\mathbf{x}) - G(\mathbf{x}) \cdot \mathbf{\Delta x} = \mathbf{0}_m,$$

где $G(\mathbf{x})$ – матрица Якоби (первых производных) для вектор-функции **g** в точке **x**. Соответственно получаем набор линейных ограничений на $\mathbf{\Delta x}$**,** *как и было*:

$$G(\mathbf{x}) \cdot \mathbf{\Delta x} = \mathbf{g}(\mathbf{x}).$$

Вектор $\mathbf{g}(\mathbf{x})$ и матрица *G*(**x**) считаются однократно при результатах измерений **x**.

Учет нелинейных ограничений в виде неравенств.

В линейной постановке мы имели ограничения вида

$$Q \cdot \mathbf{x}^* \leq \mathbf{t}$$

при $\mathbf{x}^* = \mathbf{x} - \mathbf{\Delta x}$.

Допустим, что имеем дело с нелинейным ограничением вида

$$\mathbf{r}(\mathbf{x}^*) \leq \mathbf{0}_k,$$

где **r** – тоже нелинейная вектор-функция.

Полагая погрешности и, соответственно, разности между результатами измерениями и результатами согласования малыми, т.е. $||\mathbf{\Delta x}|| \ll 1$, получим:

$$\mathbf{r}(\mathbf{x} - \mathbf{\Delta x}) = \mathbf{r}(\mathbf{x}) - R(\mathbf{x}) \cdot \mathbf{\Delta x} \leq \mathbf{0}_k,$$

где *R*(**x**) – матрица Якоби (первых производных) для вектор-функции $\mathbf{r}(\mathbf{x})$**.** Соответственно получаем набор линейных ограничений на $\mathbf{\Delta x}$**,** *как и было*:

$$(-R(\mathbf{x})) \cdot \mathbf{\Delta x} \leq -\mathbf{r}(\mathbf{x}).$$

Вектор $\mathbf{r}(\mathbf{x})$ и матрица *R*(**x**) считаются однократно при результатах измерений **x**.

Дополнение 1.

Можно попробовать учесть и матрицу Гессе, но в этом случае усложняется решение системы уравнений.

$$\mathbf{g}(\mathbf{x} - \mathbf{\Delta x}) = \mathbf{g}(\mathbf{x}) - G_1(\mathbf{x}) \cdot \mathbf{\Delta x} + \frac{1}{2} \cdot \mathbf{G_2}(\mathbf{x}) \cdot \mathbf{\Delta x}^{\otimes 2} = \mathbf{0}_m,$$

$$\mathbf{r}(\mathbf{x} - \mathbf{\Delta x}) = \mathbf{r}(\mathbf{x}) - R_1(\mathbf{x}) \cdot \mathbf{\Delta x} + \frac{1}{2} \cdot \mathbf{R_2}(\mathbf{x}) \cdot \mathbf{\Delta x}^{\otimes 2} \leq \mathbf{0}_k.$$

Здесь нижний индекс «1» в обозначения матриц указывает на матрицы Якоби (первых производных), а нижний индекс «2» – на тензор Гессе (совокупность матриц вторых производных), $\mathbf{\Delta x}^{\otimes 2}$ означает outer product.

В этом случае получим, что

$$\frac{\partial L}{\partial \mathbf{\Delta x}} = 2 \cdot \mathbf{\Delta x} + \left(\mathbf{G_2}(\mathbf{x}) \cdot \mathbf{\Delta x} - G_1(\mathbf{x})\right)^T \cdot \boldsymbol{\lambda_0} + \left(\mathbf{R_2}(\mathbf{x}) \cdot \mathbf{\Delta x} - R_1(\mathbf{x})\right)^T \cdot \boldsymbol{\lambda} = \mathbf{0}_n,$$

$$\frac{\partial L}{\partial \boldsymbol{\lambda_0}} = \mathbf{g}(\mathbf{x}) - G_1(\mathbf{x}) \cdot \mathbf{\Delta x} + \frac{1}{2} \cdot \mathbf{G_2}(\mathbf{x}) \cdot \mathbf{\Delta x}^{\otimes 2} = \mathbf{0}_m,$$

$$\frac{\partial L}{\partial \boldsymbol{\lambda}} = \mathbf{r}(\mathbf{x}) - R_1(\mathbf{x}) \cdot \mathbf{\Delta x} + \frac{1}{2} \cdot \mathbf{R_2}(\mathbf{x}) \cdot \mathbf{\Delta x}^{\otimes 2} = \mathbf{0}_k.$$

Довольно сложная система уравнений, хоть первое уравнение и линейно относительно $\mathbf{\Delta x}$. Подобную систему можно быстро решить через базисы Грёбнера.

Дополнение 2.

Если изменяется норма с $\|\mathbf{x}\|_2^2 = \mathbf{x}^T \cdot \mathbf{x}$ на $\|\mathbf{x}\|^2 = \mathbf{x}^T \cdot \mathbf{x} + \epsilon \cdot \alpha(\mathbf{x})$**,** где ϵ – заведомо малое число, а $\alpha(\mathbf{x})$ – это скалярная функция (возвращает действительное число), то получим, что первое уравнение в решаемой системе будет иметь вид

$$\frac{\partial L}{\partial \mathbf{x}} = 2 \cdot \mathbf{x} + \epsilon \cdot \boldsymbol{\alpha}'(\mathbf{x}) + A^T \cdot \boldsymbol{\lambda_0} + Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_n,$$

где $\boldsymbol{\alpha}'(\mathbf{x})$ – вектор частных производных функции $\alpha(\mathbf{x})$ по первому, второму, третьему и т.д. компоненту вектора **x** (т.е. следующих в естественном порядке).

Соответственно при подстановках из второго уравнения системы (в матричных линейных постановках), т.е. из

$$\frac{\partial L}{\partial \boldsymbol{\lambda_0}} = A \cdot \mathbf{x} - \mathbf{b} = \mathbf{0}_m,$$

получим

$$2 \cdot \mathbf{x} + \epsilon \cdot \boldsymbol{\alpha}'(\mathbf{x}) + A^T \cdot \boldsymbol{\lambda_0} + Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_n,$$

$$2 \cdot A \cdot \mathbf{x} + \epsilon \cdot A \cdot \boldsymbol{\alpha}'(\mathbf{x}) + A \cdot A^T \cdot \boldsymbol{\lambda_0} + A \cdot Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_m$$

$$2 \cdot \mathbf{b} + \epsilon \cdot A \cdot \boldsymbol{\alpha}'(\mathbf{x}) + A \cdot A^T \cdot \boldsymbol{\lambda_0} + A \cdot Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_m$$

Продолжая аналогично, получим два уравнения для $\boldsymbol{\lambda_0}$, увязывающих значение данного вектора со значениями $\boldsymbol{\lambda}$ и $\boldsymbol{\alpha}'(\mathbf{x})$. Оттуда выразим $\boldsymbol{\lambda}$ через $\boldsymbol{\alpha}'(\mathbf{x})$, а затем в итоге и $\boldsymbol{\lambda_0}$ – через $\boldsymbol{\alpha}'(\mathbf{x})$. После подстановки в первое выражение получим матричное уравнение в виде линейной формы от **x** и $\boldsymbol{\alpha}'(\mathbf{x})$. Поскольку в качестве возможных отклонений от нормальности мы используем либо полиномиальное (линейное) приближение к отличию pdf, либо первые слагаемые разложений в ряд Грама-Шарлье или Эджворта, то $\alpha(\mathbf{x})$**, безусловно,** – это полином степени не выше третьей (его коэффициенты будут зависеть от Sk и Ku распределений компонент вектора **x** и – в случае учета корреляций – его ковариационной матрицы; конкретный вид нужно выводить, используя ряд Грама-Шарлье).

В итоге получаем **полиномиальное** уравнение относительно вектора **x**, решить которого точно можно, используя матричные методы.

Случай неравноточных измерений и учет корреляций

Используем следующую норму $\|\mathbf{x}\|^2 = \mathbf{x}^T \cdot \Sigma^{-1} \cdot \mathbf{x}$, где Σ – ковариационная матрица для вектора **x** (размера nxn).

Выпишем вслед за [2] функцию Лагранжа для данной задачи:

$$L = \|\mathbf{x}\|^2 + \boldsymbol{\lambda_0} \cdot (A \cdot \mathbf{x} - \mathbf{b}) + \boldsymbol{\lambda} \cdot (Q \cdot \mathbf{x} - \mathbf{t}).$$

Необходимыми условиями достижения оптимума в соответствии с условиями Каруша-Куна-Таккера являются равенства нулю всех соответствующих частных производных. В силу симметричности матрицы Σ (коммутативности оператора ковариации), получаем, что

$$\frac{\partial L}{\partial \mathbf{x}} = 2 \cdot \Sigma^{-1} \cdot \mathbf{x} + A^T \cdot \boldsymbol{\lambda}_{\mathbf{0}} + Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_n,$$

$$\frac{\partial L}{\partial \boldsymbol{\lambda}_{\mathbf{0}}} = A \cdot \mathbf{x} - \mathbf{b} = \mathbf{0}_m,$$

$$\frac{\partial L}{\partial \boldsymbol{\lambda}} = Q \cdot \mathbf{x} - \mathbf{t} = \mathbf{0}_k.$$

Решение ищем следующим образом.

Подстановка второго уравнения в первое уравнение:

$$2 \cdot \Sigma^{-1} \cdot \mathbf{x} + A^T \cdot \boldsymbol{\lambda}_{\mathbf{0}} + Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_n$$

$$2 \cdot \Sigma \cdot \Sigma^{-1} \cdot \mathbf{x} + \Sigma \cdot A^T \cdot \boldsymbol{\lambda}_{\mathbf{0}} + \Sigma \cdot Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_n$$

$$2 \cdot \mathbf{x} + \Sigma \cdot A^T \cdot \boldsymbol{\lambda}_{\mathbf{0}} + \Sigma \cdot Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_n$$

$$2 \cdot A \cdot \mathbf{x} + A \cdot \Sigma \cdot A^T \cdot \boldsymbol{\lambda}_{\mathbf{0}} + A \cdot \Sigma \cdot Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_m$$

$$A \cdot \Sigma \cdot A^T \cdot \boldsymbol{\lambda}_{\mathbf{0}} = -2 \cdot \mathbf{b} - A \cdot \Sigma \cdot Q^T \cdot \boldsymbol{\lambda}$$

Т.к. $A \cdot \Sigma \cdot A^T$ – квадратная матрица (mxm), то

$$\boldsymbol{\lambda}_{\mathbf{0}} = -2 \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot \mathbf{b} - (A \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot \boldsymbol{\lambda}$$

Подстановка третьего уравнения в первое уравнение:

$$2 \cdot \Sigma^{-1} \cdot \mathbf{x} + A^T \cdot \boldsymbol{\lambda}_{\mathbf{0}} + Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_n$$

$$2 \cdot \Sigma \cdot \Sigma^{-1} \cdot \mathbf{x} + \Sigma \cdot A^T \cdot \boldsymbol{\lambda}_{\mathbf{0}} + \Sigma \cdot Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_n$$

$$2 \cdot \mathbf{x} + \Sigma \cdot A^T \cdot \boldsymbol{\lambda}_{\mathbf{0}} + \Sigma \cdot Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_n$$

$$2 \cdot Q \cdot \mathbf{x} + Q \cdot \Sigma \cdot A^T \cdot \boldsymbol{\lambda}_{\mathbf{0}} + Q \cdot \Sigma \cdot Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_k$$

$$Q \cdot \Sigma \cdot A^T \cdot \boldsymbol{\lambda}_{\mathbf{0}} = -2 \cdot \mathbf{t} - Q \cdot \Sigma \cdot Q^T \cdot \boldsymbol{\lambda}$$

Т.к. $A \cdot \Sigma \cdot A^T$ – неквадратная матрица (kxm), то

$$(Q \cdot \Sigma \cdot A^T)^T \cdot Q \cdot \Sigma \cdot A^T \cdot \boldsymbol{\lambda}_{\mathbf{0}} = -2 \cdot (Q \cdot \Sigma \cdot A^T)^T \cdot \mathbf{t} - (Q \cdot \Sigma \cdot A^T)^T \cdot Q \cdot \Sigma \cdot Q^T \cdot \boldsymbol{\lambda}$$

Т.к. $(Q \cdot \Sigma \cdot A^T)^T \cdot Q \cdot \Sigma \cdot A^T = A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T$, то

$$A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T \cdot \boldsymbol{\lambda}_{\mathbf{0}} = -2 \cdot A \cdot \Sigma \cdot Q^T \cdot \mathbf{t} - A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot Q^T \cdot \boldsymbol{\lambda}$$

$$\boldsymbol{\lambda}_{\mathbf{0}} = -2 \cdot (A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot \mathbf{t} - (A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot Q^T \cdot \boldsymbol{\lambda}$$

Приравняем полученные равенства:

$$-2 \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot \mathbf{b} - (A \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot \boldsymbol{\lambda} = -2 \cdot (A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot \mathbf{t} - (A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot Q^T \cdot \boldsymbol{\lambda}$$

$$((A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot Q^T - (A \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T) \cdot \boldsymbol{\lambda} = -2 \cdot (A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot \mathbf{t} - 2 \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot \mathbf{b}$$

Матрица $P = (A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot Q^T - (A \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T$ не является квадратной (mxk), поэтому

$$P^T \cdot P \cdot \boldsymbol{\lambda} = -2 \cdot P^T \cdot (A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot \mathbf{t} - 2 \cdot P^T \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot \mathbf{b}$$

Откуда

$$\boldsymbol{\lambda} = -2 \cdot (P^T \cdot P)^{-1} \cdot P^T \cdot (A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot \mathbf{t} - 2 \cdot (P^T \cdot P)^{-1} \cdot P^T \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot \mathbf{b}$$

Следовательно,

$$\boldsymbol{\lambda}_\mathbf{0} = -2 \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot \mathbf{b} - (A \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot \boldsymbol{\lambda} =$$

$$-2 \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot \mathbf{b} + 2 \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot (P^T \cdot P)^{-1} \cdot P^T \cdot (A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot \mathbf{t} + 2 \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot (P^T \cdot P)^{-1} \cdot P^T \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot \mathbf{b} =$$

$$2 \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot (A \cdot \Sigma \cdot Q^T \cdot (P^T \cdot P)^{-1} \cdot P^T \cdot (A \cdot \Sigma \cdot A^T)^{-1} - E_m) \cdot \mathbf{b} + 2 \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot (P^T \cdot P)^{-1} \cdot P^T \cdot (A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot \mathbf{t}$$

Наконец,

$$2 \cdot \mathbf{x} + \Sigma \cdot A^T \cdot \boldsymbol{\lambda}_\mathbf{0} + \Sigma \cdot Q^T \cdot \boldsymbol{\lambda} = \mathbf{0}_n$$

$$\mathbf{x} = -\Sigma \cdot A^T \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot (A \cdot \Sigma \cdot Q^T \cdot (P^T \cdot P)^{-1} \cdot P^T \cdot (A \cdot \Sigma \cdot A^T)^{-1} - E_m) \cdot \mathbf{b} - \Sigma \cdot A^T \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot (P^T \cdot P)^{-1} \cdot P^T \cdot (A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot \mathbf{t} + \Sigma \cdot Q^T \cdot (P^T \cdot P)^{-1} \cdot P^T \cdot (A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot \mathbf{t} + \Sigma \cdot Q^T \cdot (P^T \cdot P)^{-1} \cdot P^T \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot \mathbf{b} = W_b \cdot \mathbf{b} + W_t \cdot \mathbf{t},$$

где $W_b = \Sigma \cdot Q^T \cdot (P^T \cdot P)^{-1} \cdot P^T \cdot (A \cdot \Sigma \cdot A^T)^{-1} - \Sigma \cdot A^T \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot (A \cdot \Sigma \cdot Q^T \cdot (P^T \cdot P)^{-1} \cdot P^T \cdot (A \cdot \Sigma \cdot A^T)^{-1} - E_m)$,

$$W_t = \Sigma \cdot Q^T \cdot (P^T \cdot P)^{-1} \cdot P^T \cdot (A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T - \Sigma \cdot A^T \cdot (A \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T \cdot (P^T \cdot P)^{-1} \cdot P^T \cdot (A \cdot \Sigma \cdot Q^T \cdot Q \cdot \Sigma \cdot A^T)^{-1} \cdot A \cdot \Sigma \cdot Q^T.$$

Определение вида функции $\alpha(\mathbf{x})$

Используем следующую норму $||\mathbf{\Delta x}||^2 = \mathbf{\Delta x}^T \cdot \Sigma^{-1} \cdot \mathbf{\Delta x} + \epsilon \cdot \alpha(\mathbf{\Delta x})$, где Σ – ковариационная матрица для вектора **x** (размера nxn). Предположим следующее: 1) всеми внедиагональными кумулянтами высоких порядков можно пренебречь (то есть оперировать только коэффициентами асимметрии *Sk* и эксцессом *Ex* каждой компоненты вектора **x** отдельно); 2) отличия от нормальности малы, т.е. |*Sk*|<<1 и $|\widehat{Ku}|$=|*Ku*–3|<<1. Тогда метод максимального правдоподобия приводит к метрике вида

$$||\mathbf{\Delta x}||^2 = \mathbf{\Delta x}^T \cdot \Sigma^{-1} \cdot \mathbf{\Delta x} + \mathbf{\Delta x}^T \cdot Q^T \cdot \left(\frac{1}{3} \cdot Q \cdot \mathbf{\Delta x} \cdot \mathbf{\Delta x}^T \cdot Q^T - E_n\right) \cdot \mathbf{Sk} + \cdot \left(\frac{1}{12} \cdot (Q \cdot \mathbf{\Delta x} \cdot \mathbf{\Delta x}^T \cdot Q^T)^2 - \frac{1}{2} \cdot Q \cdot \mathbf{\Delta x} \cdot \mathbf{\Delta x}^T \cdot Q^T + \frac{1}{4} \cdot E_n\right) \cdot \widehat{\mathbf{Ku}}$$

где $Q^T \cdot Q = \Sigma^{-1}$ – разложение Холецкого для матрицы Σ^{-1}. Возведение в квадрат в формуле понимается в матричном смысле, поскольку матрица $Q \cdot \mathbf{\Delta x} \cdot \mathbf{\Delta x}^T \cdot Q^T$ является квадратной и симметричной: $(Q \cdot \mathbf{\Delta x} \cdot \mathbf{\Delta x}^T \cdot Q^T)^T = Q \cdot \mathbf{\Delta x} \cdot \mathbf{\Delta x}^T \cdot Q^T$.