

СТРУКТУРА КУРСА

Теория

пакет Statistica

Приложения

Корреляционный анализ

Регрессионный анализ

Простая

Множественная

Нелинейная

Дисперсионный анализ

Многомерный анализ

Факторный

Кластерный

Дискриминантный

Планирование эксперимента

Оптимизация планов

Последовательное планирование

Стохастическая аппроксимация

Статистика автокомбината

Центробанк

АСМАП

Маршрутная сеть

Террамеханика

Характеристика автомобилей

Поставки запчастей

Корреляционный анализ

Простая регрессия

Множественная регрессия

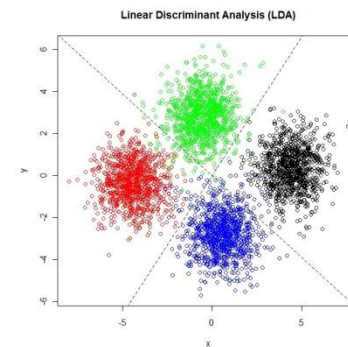
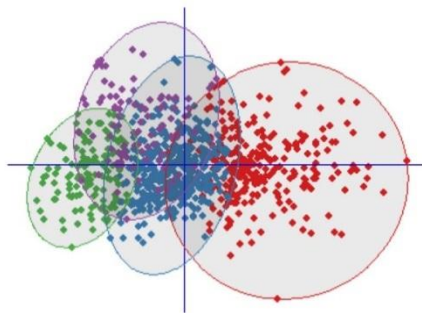
Нелинейная регрессия

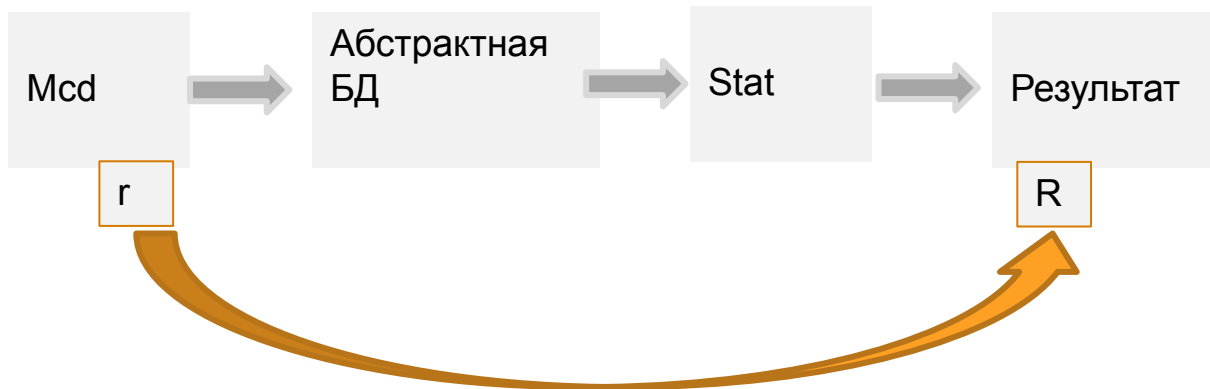
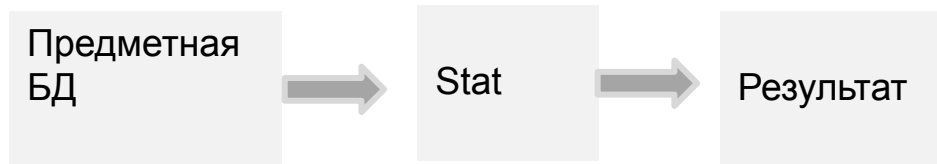
$$\sum_{i=1}^n d_i^2 \rightarrow \mathit{min}$$

Дисперсионный анализ

Кластерный анализ

Дискриминантный анализ





matrix(L,N,f) - создание матрицы - **L** – число строк, **N** – число столбцов, **f** - функция **f(l,n)**

identity(n) - создание единичной матрицы размерности **n**

diag(v) - создание диагональной матрицы с элементами главной диагонали из **v**;

geninv(A) - обращение матрицы **A**, аналогично операции **A⁻¹**.

rows(A) - определение числа строк матрицы **A**

cols(A) - определение числа столбцов матрицы **A**

last(v) - вычисление номера последнего элемента вектора **v**

length(v) - вычисление количества элементов **v**

csort(A,i), rsort(A,j) - сортировка по возрастанию элементов **i**-го столбца или строки

min(A), max(A) - минимальный и максимальный элемент матрицы

tr(A) - след квадратной матрицы **A**

rank(A) - ранг матрицы **A**

norm1(A), norm2(A), norme(A), normi(A) - нормы квадратной матрицы **A**

submatrix (M, r1, r2, c1, c2) - подматрица **M**, **r1** и **r2** – нижний и верхний номер строки, **c1** и **c2** - нижней и верхний номер столбца матрицы **M**

augment(A,B,...), stack(A,B,...) - слияние матриц **A, B, ...** слева направо или сверху вниз

rref(A) - приведение матрицы к ступенчатому виду с единичным базисным минором

eigenvals(A) - собственные значения квадратной матрицы **A**

eigenvecs(A) - собственные вектора квадратной матрицы **A** по порядку собственных значений

eigenvec(A, l) - собственный вектор для собственного значения

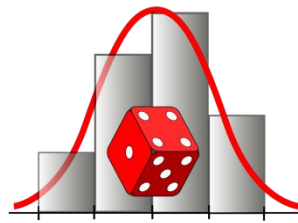
lsolve(A, b) - решение СЛАУ **Ax=b**

Обозначение	Теория множеств	Теория вероятностей
Ω	Множеств, пространство	Пространство элементарных событий
ω	Элемент множества	Элементарное событие
A, B	Подмножество A, B	Случайное событие A, B
$A+B = A \cup B$	Объединение множеств	Сумма событий
AB	Пересечение множеств	Произведение событий
\bar{A}	Дополнение множества	Противоположное событие
A/B	Разность множеств	Разность событий
\emptyset	Пустое множество	Невозможное событие
$AB = \emptyset$	Множества не пересекаются	События несовместны
$A=B$	Множества равны	События равносильны
$A \subset B$	A есть подмножество B	Событие A влечет событие B



Пространство элементарных событий

- Любая природа
- Не обязательно числовая



Случайная величина –
числовая функция на пространстве
элементарных событий



Числовые характеристики случайных величин

Математическое ожидание - $\sum x_i p_i$

Дисперсия - $D\xi = M(\xi - M\xi)^2$

Ковариация - $cov(\xi, \eta) = M(\xi - M\xi)(\eta - M\eta)$

Коэффициент корреляции $\rho_{\xi, \eta} = \frac{cov(\xi, \eta)}{\sqrt{D\xi \cdot D\eta}}$

Свойства математического ожидания

Свойства дисперсии и ковариации

- $\xi \geq 0 \Rightarrow M\xi \geq 0$;
- $M a\xi = aM\xi$
- $M(\xi + \eta) = M\xi + M\eta$;
- $M(\xi\eta) = M\xi M\eta$ - для независимых случайных величин
- $M(|\xi\eta|)^2 \leq M\xi M\eta$ - Неравенство Коши-Буняковского

- $D\xi = M\xi^2 - (M\xi)^2$,
- $D(a+b\xi) = b^2 D\xi$
- $cov(\xi, \eta) = M(\xi\eta) - M(\xi)M(\eta)$

Случайный вектор

$$\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \dots \\ \xi_n \end{bmatrix} \quad \xi = \begin{bmatrix} \xi_1 & \xi_2 & \dots & \xi_n \end{bmatrix}$$

Математическое ожидание

$$M \xi = \begin{bmatrix} M \xi_1 \\ M \xi_2 \\ \dots \\ M \xi_n \end{bmatrix}$$

$$M \xi = \begin{bmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{bmatrix}$$

Дисперсионная матрица ξ - $D\xi = \|\text{cov}(\xi_i, \xi_j)\|_{ij=1..n}$

Свойства $D\xi$: $D\xi$ – симметрична ($D=D^T$)
 $D\xi$ - неотрицательно-определенная ($\forall X: X^T D X \geq 0$)

D - дисперсионная матрица $\Rightarrow D=A^T A$

Представление D : $D\xi = M[(\xi - M\xi)(\xi - M\xi)^T]$

Некоторые доказательства

$$\sum_{i,j} \text{cov}(\xi_i, \xi_j) x_i x_j = \sum_{i,j} M(\xi_i - M\xi_i)(\xi_j - M\xi_j) \cdot x_i \cdot x_j = M \sum_{i,j} [(\xi_i - M\xi_i) \cdot x_i]^2 \geq 0$$

$$D=A^T A \Rightarrow D^T = (A^T A)^T = A^T (A^T)^T = A^T A = D$$

$$D=A^T A \Rightarrow \forall X: X^T D X = X^T (A^T A) X = (X^T A^T)(A X) = (A X)^T (A X) = Y^T Y = \sum y_i^2 \geq 0$$

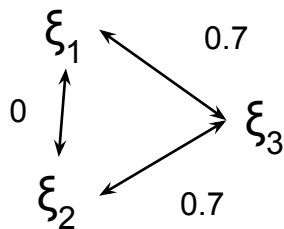
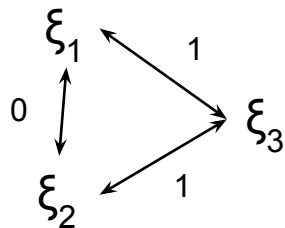
$$\begin{aligned}
 \mathbf{M}[(\xi - \mathbf{M}\xi)(\xi - \mathbf{M}\xi)^T] &= \begin{bmatrix} \xi_1 - \mathbf{M}\xi_1 \\ \dots \\ \xi_i - \mathbf{M}\xi_i \\ \dots \\ \xi_n - \mathbf{M}\xi_n \end{bmatrix} \begin{bmatrix} \xi_1 - \mathbf{M}\xi_1 & \dots & \xi_j - \mathbf{M}\xi_j & \dots & \xi_n - \mathbf{M}\xi_n \end{bmatrix} = \\
 &= \mathbf{M} \begin{bmatrix} (\xi_1 - \mathbf{M}\xi_1)^2 & \dots & (\xi_1 - \mathbf{M}\xi_1)(\xi_j - \mathbf{M}\xi_j) & \dots & (\xi_1 - \mathbf{M}\xi_1)(\xi_n - \mathbf{M}\xi_n) \\ \dots & \dots & \dots & \dots & \dots \\ (\xi_i - \mathbf{M}\xi_i)(\xi_1 - \mathbf{M}\xi_1) & \dots & (\xi_i - \mathbf{M}\xi_i)(\xi_j - \mathbf{M}\xi_j) & \dots & (\xi_i - \mathbf{M}\xi_i)(\xi_n - \mathbf{M}\xi_n) \\ \dots & \dots & \dots & \dots & \dots \\ (\xi_n - \mathbf{M}\xi_n)(\xi_1 - \mathbf{M}\xi_1) & \dots & (\xi_n - \mathbf{M}\xi_n)(\xi_j - \mathbf{M}\xi_j) & \dots & (\xi_n - \mathbf{M}\xi_n)^2 \end{bmatrix} = \\
 &= \begin{bmatrix} \mathbf{M}(\xi_1 - \mathbf{M}\xi_1)^2 & \dots & \mathbf{M}(\xi_1 - \mathbf{M}\xi_1)(\xi_j - \mathbf{M}\xi_j) & \dots & \mathbf{M}(\xi_1 - \mathbf{M}\xi_1)(\xi_n - \mathbf{M}\xi_n) \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{M}(\xi_i - \mathbf{M}\xi_i)(\xi_1 - \mathbf{M}\xi_1) & \dots & \mathbf{M}(\xi_i - \mathbf{M}\xi_i)(\xi_j - \mathbf{M}\xi_j) & \dots & \mathbf{M}(\xi_i - \mathbf{M}\xi_i)(\xi_n - \mathbf{M}\xi_n) \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{M}(\xi_n - \mathbf{M}\xi_n)(\xi_1 - \mathbf{M}\xi_1) & \dots & \mathbf{M}(\xi_n - \mathbf{M}\xi_n)(\xi_j - \mathbf{M}\xi_j) & \dots & \mathbf{M}(\xi_n - \mathbf{M}\xi_n)^2 \end{bmatrix} = \|\text{cov}(\xi_i, \xi_j)\| = \mathbf{D}\xi
 \end{aligned}$$

$$C=A*B \quad c_{ij} = \sum_{k=1} a_{ik} b_{kj}$$

$$(A*B)^T = B^T*A^T$$

Линейное преобразование математического ожидания случайного вектора равно математическому ожиданию линейного преобразования B ($M\xi) = M(B\xi)$

Взаимные корреляции



Критерий
неотрицательно
определенной матрицы

Задан вектор ξ

$$\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \dots \\ \xi_n \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

ξ - случайный вектор
 \mathbf{A} и \mathbf{B} – числовые матрицы
 $D_\xi = \mathbf{D}\xi = \|\text{cov}(\xi_i, \xi_j)\|_{i,j=1..n}$ - дисперсионная матрица ξ
 $M_\xi = \mathbf{M}\xi$ - математическое ожидание ξ

Случайный вектор η представляет линейное преобразование вектора ξ

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \dots \\ \eta_m \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \dots \\ \xi_n \end{bmatrix} \longrightarrow \eta = \mathbf{A} + \mathbf{B}\xi \longrightarrow$$

Задача: найти
 M_η - математическое ожидание η
 $D_{\eta\eta} = \mathbf{D}\eta = \|\text{cov}(\eta_i, \eta_j)\|_{i,j=1..m}$

Утверждение

$$\mathbf{M}_\eta = \mathbf{A} + \mathbf{B}\mathbf{M}_\xi \quad \mathbf{D}_\eta = \mathbf{B}\mathbf{D}_\xi\mathbf{B}^T$$

$$\mathbf{M}\eta = \mathbf{M}(\mathbf{A} + \mathbf{B}\xi) = \mathbf{M}\mathbf{A} + \mathbf{M}(\mathbf{B}\xi) = \mathbf{A} + \mathbf{B}(\mathbf{M}\xi)$$

$$\begin{aligned} \mathbf{D}\eta &= \mathbf{M}[(\eta - \mathbf{M}\eta)(\eta - \mathbf{M}\eta)^T] = \mathbf{M}[(\mathbf{A} + \mathbf{B}\xi - \mathbf{M}(\mathbf{A} + \mathbf{B}\xi))(\mathbf{A} + \mathbf{B}\xi - \mathbf{M}(\mathbf{A} + \mathbf{B}\xi))^T] = \mathbf{M}[(\mathbf{B}\xi - \mathbf{M}\mathbf{B}\xi)(\mathbf{B}\xi - \mathbf{M}\mathbf{B}\xi)^T] = \\ &= \mathbf{M}[\mathbf{B}(\xi - \mathbf{M}\xi)(\xi - \mathbf{M}\xi)^T\mathbf{B}^T] = \mathbf{B}\mathbf{M}[(\xi - \mathbf{M}\xi)(\xi - \mathbf{M}\xi)^T]\mathbf{B}^T = \mathbf{B}\mathbf{D}_\xi\mathbf{B}^T. \end{aligned}$$

Оптимальное оценивание нормально-распределенных случайных векторов

$$\xi = (\xi_1, \dots, \xi_m)^T \longrightarrow \eta = (\eta_1, \dots, \eta_m)^T$$

Минимизация суммы квадратов

Совместный случайный вектор $(\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_m)^T$

$$M \begin{bmatrix} \xi \\ \eta \end{bmatrix} = \begin{bmatrix} m_\xi \\ m_\eta \end{bmatrix} \quad - \text{математическое ожидание совместного вектора}$$

$$D \begin{bmatrix} \xi \\ \eta \end{bmatrix} = \begin{bmatrix} D_{\xi\xi} & D_{\xi\eta} \\ D_{\eta\xi} & D_{\eta\eta} \end{bmatrix} \quad - \text{дисперсионная матрица совместного вектора}$$

$$D_{\eta\eta} \equiv \|\text{cov}(\eta_i, \eta_j)\| \quad 1 \leq i, j \leq m$$

$$D_{\eta\xi} \equiv \|\text{cov}(\eta_i, \xi_j)\| \quad 1 \leq i \leq m, \quad 1 \leq j \leq n$$

$$D_{\xi\xi} \equiv \|\text{cov}(\xi_i, \xi_j)\| \quad 1 \leq i, j \leq n$$

Утверждение

Оценка вектора η

$$M(\eta | \xi) = m_\eta + D_{\eta\xi} D_{\xi\xi}^{-1} (\xi - m_\xi)$$

Ошибка оценка вектора η

$$\Delta = D_{\eta\eta} - D_{\eta\xi} D_{\xi\xi}^{-1} D_{\xi\eta}$$

$$\zeta(\xi, \eta) = (\eta - m_\eta) - D_{\eta\xi} D_{\xi\xi}^{-1} (\xi - m_\xi)$$

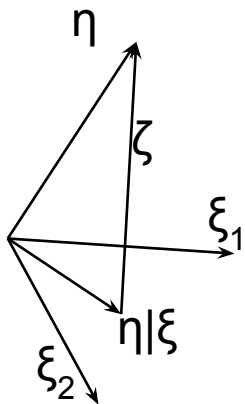
$$M\zeta = M(\eta - m_\eta) - M[D_{\eta\xi} D_{\xi\xi}^{-1} (\xi - m_\xi)] = 0$$

$$D_{\zeta\xi} = M[(\zeta - m_\zeta)(\xi - m_\xi)^T] = 0$$

Доказать

Вывод: ξ и ζ некоррелированы \Rightarrow в силу нормальности распределения - независимы

Оптимальная оценка ζ по ξ это $M\zeta=0$ - 1-ое утверждение доказано



$$D_{\zeta\xi} = M[\zeta\xi^T] = D_{\eta\xi} - D_{\eta\xi} D_{\xi\xi}^{-1} D_{\xi\eta}$$

Доказать

$$\begin{aligned}
 D_{\zeta\xi} &= M[(\zeta - m_\zeta)(\xi - m_\xi)^T] = M[\zeta(\xi - m_\xi)] = M[(\eta - m_\eta) - D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)^T](\xi - m_\xi)] = \\
 &= M[(\eta - m_\eta)(\xi - m_\xi)] - D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)^T(\xi - m_\xi)] = M[(\eta - m_\eta)(\xi - m_\xi)] - D_{\eta\xi} D_{\xi\xi}^{-1}[M(\xi - m_\xi)^T(\xi - m_\xi)] = \\
 &= D_{\eta\xi} - D_{\eta\xi} D_{\xi\xi}^{-1} D_{\xi\xi} = 0
 \end{aligned}$$

$$\begin{aligned}
 D_{\zeta\zeta} &= M[\zeta\zeta^T] = [(\eta - m_\eta) - D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)][(\eta - m_\eta) - D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)]^T = \\
 &= M \left[\begin{aligned} &(\eta - m_\eta)(\eta - m_\eta)^T - (\eta - m_\eta)[D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)]^T - \\ &- D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)(\eta - m_\eta)^T + D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)[D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)]^T \end{aligned} \right] = \\
 &= M[(\eta - m_\eta)(\eta - m_\eta)^T] - M(\eta - m_\eta)(\xi - m_\xi)^T D_{\xi\xi}^{-1} D_{\xi\eta} - \\
 &- MD_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)(\eta - m_\eta)^T + MD_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)(\xi - m_\xi)^T D_{\xi\xi}^{-1} D_{\xi\eta} \\
 &= D_{\eta\eta} - D_{\eta\xi} D_{\xi\xi}^{-1} D_{\xi\eta} - D_{\eta\xi} D_{\xi\xi}^{-1} D_{\xi\eta} + D_{\eta\xi} D_{\xi\xi}^{-1} D_{\xi\xi} D_{\xi\xi}^{-1} D_{\xi\eta}^T = \\
 &= D_{\eta\eta} - D_{\eta\xi} D_{\xi\xi}^{-1} D_{\xi\eta}
 \end{aligned}$$

СЛУЧАЙ НЕЗАВИСИМЫХ ПРЕДИКТОРОВ И СКАЛЯРНОЙ ПРОГНОЗИРУЕМОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

$(\xi_1, \dots, \xi_n, \eta)^T$

$(n+1)$ -мерный гауссовский вектор, причем $\xi_1, \xi_2 \dots \xi_n$ попарно независимы $\text{cov}(\xi_i, \xi_j)=0$, а η является одномерной величиной,

 определение размеров матрицы

X_n - ввод нижнего индекса

X^{-1} - вычисление обратной матрицы

$|X|$ - определитель матрицы $|M| = \det M$ или длина вектора $|x|$

→ - поэлементные операции с матрицами

$M^{<j>}$ - $M^{<j>}$ j-й столбец матрицы

M^T - транспонирование матрицы

$\hat{u} \cdot \hat{e}$ - скалярное произведение векторов $x \cdot y = \sum x_i y_i$

$\hat{u} \times \hat{e}$ - векторное произведение $a \times b = (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1)$;

$\sum X$ - сумма компонент вектора $\sum x = \sum x_i$

m..n - диапазон изменения переменной

$$\begin{aligned}
 D_{\zeta\xi} &= M[(\zeta - m_\zeta)(\xi - m_\xi)^T] = M[\zeta(\xi - m_\xi)] = \\
 &= M[(\eta - m_\eta) - D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)^T](\xi - m_\xi) = \\
 &= M[(\eta - m_\eta)(\xi - m_\xi)] - D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)^T(\xi - m_\xi) = \\
 &= M[(\eta - m_\eta)(\xi - m_\xi)] - D_{\eta\xi} D_{\xi\xi}^{-1}[M(\xi - m_\xi)^T(\xi - m_\xi)] = \\
 &= D_{\eta\xi} - D_{\eta\xi} D_{\xi\xi}^{-1} D_{\xi\xi}
 \end{aligned}$$

$$\begin{aligned}
 D_{\zeta\zeta} &= M[\zeta\zeta^T] = [(\eta - m_\eta) - D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)][(\eta - m_\eta) - D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)]^T = \\
 &= M \left[\begin{array}{l} (\eta - m_\eta)(\eta - m_\eta)^T - (\eta - m_\eta)[D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)]^T - \\ - D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)(\eta - m_\eta)^T + D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)[D_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)]^T \end{array} \right] = \\
 &= M[(\eta - m_\eta)(\eta - m_\eta)^T] - M(\eta - m_\eta)(\xi - m_\xi)^T D_{\xi\xi}^{-1} D_{\xi\eta} - \\
 &\quad - MD_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)(\eta - m_\eta)^T + MD_{\eta\xi} D_{\xi\xi}^{-1}(\xi - m_\xi)(\xi - m_\xi)^T D_{\xi\xi}^{-1} D_{\xi\eta} \\
 &= D_{\eta\eta} - D_{\eta\xi} D_{\xi\xi}^{-1} D_{\xi\eta} - D_{\eta\xi} D_{\xi\xi}^{-1} D_{\xi\eta} + D_{\eta\xi} D_{\xi\xi}^{-1} D_{\xi\xi} D_{\xi\xi}^{-1} D_{\xi\xi}^T = \\
 &= D_{\eta\eta} - D_{\eta\xi} D_{\xi\xi}^{-1} D_{\xi\eta}^T
 \end{aligned}$$