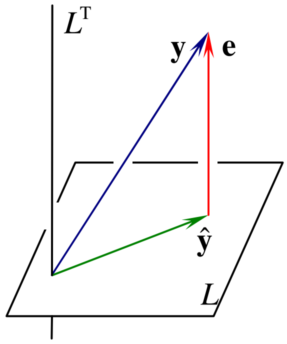
Постановка задачи

Пусть L – линейное подпространство евклидова пространства E. Требуется представить вектор $\mathbf{y}$ в виде суммы $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{e}$, где $\hat{\mathbf{y}} \in L$, $\mathbf{e} \in L^{\mathrm{T}}$, т.е. вектор $\hat{\mathbf{y}}$ принадлежит подпространству L, а вектор $\mathbf{e}$ ортогонален L.



По теореме о разложении евклидова пространства, для любого подпространства $L \subseteq E$ имеет место разложение $E = L \oplus L^{\mathrm{T}}$, поэтому представление $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{e}$ всегда существует и единственно.

Векторы $\hat{\mathbf{y}}$ и $\mathbf{e}$ называются *ортогональной проекцией вектора* $\mathbf{y}$ *на подпространство* L и *ортогональной составляющей* соответственно и обозначаются $\hat{\mathbf{y}} = \mathrm{Pr}_L(\mathbf{y})$ и $\mathbf{e} = \mathrm{Ort}_L(\mathbf{y})$.

Замечания

1. Так как векторы $\hat{\mathbf{y}}$ и $\mathbf{e}$ ортогональны, то по теореме Пифагора $\|\mathbf{y}\|^2 = \|\hat{\mathbf{y}}\|^2 + \|\mathbf{e}\|^2$.
2. Если $\mathbf{y} \in L$, то $\hat{\mathbf{y}} = \mathbf{y}$ и $\mathbf{e} = \mathbf{O}$; если $\mathbf{y} \in L^{\mathrm{T}}$, то $\hat{\mathbf{y}} = \mathbf{O}$ и $\mathbf{e} = \mathbf{y}$.

Экстремальное свойство ортогональной проекции вектора

Пусть в линейном подпространстве L евклидова пространства E требуется найти вектор $\mathbf{z}$ ближайший к вектору $\mathbf{y}$. Для квадрата расстояния между этими векторами имеем следующую цепочку равенств:

$$\rho^2(\mathbf{y},\mathbf{z}) = \|\mathbf{y} - \mathbf{z}\|^2 = \|\hat{\mathbf{y}} + \mathbf{e} - \mathbf{z}\|^2 = \|\hat{\mathbf{y}} - \mathbf{z} + \mathbf{e}\|^2 .$$

Так как разность векторов $\hat{\mathbf{y}} - \mathbf{z}$ принадлежат подпространству L, а вектор $\mathbf{e}$ ортогонален подпространству L, то по теореме Пифагора следует, что

$$\rho^2(\mathbf{y},\mathbf{z}) = \|\hat{\mathbf{y}} - \mathbf{z}\|^2 + \|\mathbf{e}\|^2 .$$

Из полученного соотношения следует, что квадрат расстояния между векторами $\mathbf{y}$ и $\mathbf{z}$ представим в виде суммы двух неотрицательных слагаемых и будет минимальным, если вектор $\mathbf{z} = \hat{\mathbf{y}}$. Таким образом, ближайшим к вектору $\mathbf{y}$ в подпространстве L является вектор $\hat{\mathbf{y}} = \mathrm{Pr}_L(\mathbf{y})$.

Система нормальных уравнений

Пусть в подпространстве L имеется базис $\{\mathbf{x}\} = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k\}$, тогда вектор $\hat{\mathbf{y}}$ может быть представлен в виде $\hat{\mathbf{y}} = b_1\mathbf{x}_1 + b_2\mathbf{x}_2 + ... + b_k\mathbf{x}_k$, где величины b_1, b_2,…, b_k – координаты вектора $\hat{\mathbf{y}}$ в базисе $\{\mathbf{x}\}$. Таким образом, задача проектирования вектора $\mathbf{y}$ на линейное подпространство L сводится к задаче определения координат вектора $\hat{\mathbf{y}}$ в базисе $\{\mathbf{x}\}$.

По условию задачи $\mathbf{e} \in L^{\mathrm{T}}$, это значит, что вектор $\mathbf{e}$ ортогонален любому вектору из L, в том числе любому из векторов базиса $\{\mathbf{x}\}$, поэтому $(\mathbf{x}_i, \mathbf{e}) = 0$, для $i = 1,2,...,k$.

Из равенства $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{e}$ следует, что $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$, поэтому

$$(\mathbf{x}_i, \mathbf{e}) = (\mathbf{x}_i, \mathbf{y} - \hat{\mathbf{y}}) = (\mathbf{x}_i, \mathbf{y}) - (\mathbf{x}_i, \hat{\mathbf{y}}) = 0, \quad i = 1,2,...,k$$

или

$$(\mathbf{x}_i, \hat{\mathbf{y}}) = (\mathbf{x}_i, \mathbf{y}), \quad i = 1,2,...,k .$$

Заменив в последнем выражении $\hat{\mathbf{y}}$ его разложением по векторам базиса $\{\mathbf{x}\}$, получим *систему нормальных уравнений*

$$\sum_{j=1}^{k} b_j (\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i, \mathbf{y}), \quad i = 1,2,...,k .$$

Матрица системы нормальных уравнений

$$\mathbf{\Gamma} = \begin{bmatrix} (\mathbf{x}_1,\mathbf{x}_1) & (\mathbf{x}_1,\mathbf{x}_2) & \dots & (\mathbf{x}_1,\mathbf{x}_k) \\ (\mathbf{x}_2,\mathbf{x}_1) & (\mathbf{x}_2,\mathbf{x}_2) & \dots & (\mathbf{x}_2,\mathbf{x}_k) \\ \dots & \dots & \dots & \dots \\ (\mathbf{x}_k,\mathbf{x}_1) & (\mathbf{x}_k,\mathbf{x}_2) & \dots & (\mathbf{x}_k,\mathbf{x}_k) \end{bmatrix}$$

называется *матрицей Грама*. Если её определитель не равен нулю, то решение системы может быть найдена методом Крамера:

$$b_j = \frac{\Delta_j}{|\mathbf{\Gamma}|},\ j = 1,2,\dots,k\ ,$$

где Δ_j – это определитель матрицы, получающейся из матрицы Грама путём замены j-го столбца на столбец свободных членов системы нормальных уравнений.

Система нормальных уравнений в матричном виде

Пусть в роли линейного пространства E выступает пространство R_n со стандартным скалярным произведением $(\mathbf{x},\mathbf{y}) = \mathbf{x}^{\mathrm{T}}\mathbf{y}$, тогда векторы $\mathbf{y}$, $\hat{\mathbf{y}}$ и $\mathbf{x}_1$, $\mathbf{x}_2,\dots,\mathbf{x}_k$ имеют вид:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},\ \hat{\mathbf{y}} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix} \text{ и } \mathbf{x}_j = \begin{bmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{bmatrix},\ j = 1,2,\dots,k\ .$$

Введём обозначения:

1. $\mathbf{X} = \left[\mathbf{x}_1 \,|\, \mathbf{x}_2 \,|\cdots|\, \mathbf{x}_k\right]$ – матрица размера $n \times k$, j-й столбец которой представляет собой вектор-столбец $\mathbf{x}_j$.
2. $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}$ – координатный вектор-столбец вектора $\hat{\mathbf{y}}$ в базисе $\{\mathbf{x}\} = \{\mathbf{x}_1,\mathbf{x}_2,\dots,\mathbf{x}_k\}$.

С учётом введённых обозначений система нормальных уравнений примет вид:

$$(\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{b} = \mathbf{X}^{\mathrm{T}}\mathbf{y}\ .$$

Если матрица $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ является невырожденной, то она имеет обратную матрицу и система нормальных уравнений будет иметь решение:

$$\mathbf{b} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}\ ,$$

в этом случае искомый вектор $\hat{\mathbf{y}}$ будет равен

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{b} \quad \text{или} \quad \hat{\mathbf{y}} = \mathbf{P}\mathbf{y}\ ,$$

где матрица $\mathbf{P} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$ называется *матрицей ортогонального проектирования на подпространство* L.

Из теоремы (о совместности системы $\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{A}^{\mathrm{T}}\mathbf{b}$) следует, что система нормальных уравнений всегда совместна. Интересует вопрос, при каких условиях данная система имеет единственное решение? Ответ на этот вопрос даёт следующая теорема.

Теорема (о решении системы нормальных уравнений)

Система нормальных уравнений всегда совместна и имеет единственное решение тогда и только тогда, когда $r(\mathbf{X}) = k$, т.е. когда ранг матрицы $\mathbf{X}$ равен количеству её столбцов.