Linearna regresija

Imamo N točk, za katere želimo najti elipso ki se tem točkam najbolj prilega.

$$T_1(x_1, y_1), T_2(x_2, y_2) \ldots, T_n(x_n, y_n)$$

Zapišemo splošno enačbo za krivulje 2. reda:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

Ker je točk preveč, ne moremo najti elipse, na kateri bodo ležale vse tocke. Zato pokusamo najti stožnico (elipso), ki se točkam najbolje prilega. Za točko T_i velja:

$$\varepsilon_i = Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F$$

Pri čemer je ε_i napaka v tej tocki. Iščemo najboljše parametre A, B, C, D, E in F, tako da bo skupna napaka čim manjša.

Ko izračunamo napako za vsako točko dobimo vektor napak. Zanima nas skupna velikost napake, kar je 'dolžina' tega vektorja.

$$\varepsilon = \sqrt{\sum_{i=1}^{N}(Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F)^2}$$

Želimo najti minimum te funkcije, ki je odvisna od 6-ih spremenljivk. Če iščemo minimum te funkcije, je enako kot da bi iskali minimum ε^2.

$$\varepsilon^2(A, B, C, D, E, F) = \sum_{i=1}^{N}(Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F)^2$$

Izračunamo odvode za to funkcijo, po vseh spremenljivkah.

$$\frac{\partial\varepsilon}{\partial A} = \sum_{i=1}^{N} 2(Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F)(x_i^2)$$

$$\frac{\partial\varepsilon}{\partial B} = \sum_{i=1}^{N} 2(Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F)(x_iy_i)$$

$$\frac{\partial \varepsilon}{\partial C} = \sum_{i=1}^{N} 2(Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F)(y_i^2)$$

$$\frac{\partial \varepsilon}{\partial D} = \sum_{i=1}^{N} 2(Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F)(x_i)$$

$$\frac{\partial \varepsilon}{\partial E} = \sum_{i=1}^{N} 2(Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F)(y_i)$$

$$\frac{\partial \varepsilon}{\partial F} = \sum_{i=1}^{N} 2(Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F)$$

Da najdemo rešitev, moramo ugotoviti kdaj so vsi odvodi enaki 0. Tega ne znamo (možno je z uporabo matrik, ampak ne znamo), zato se bomo reševanja lotili z gradientnim spustom.

Gradientni spust

Gradient nam pove smer največjega naraščanja funkcije. Gradient funkcije f(x) je definiran kot:

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \frac{\partial f}{\partial x_3} & \dots & \frac{\partial f}{\partial x_n} \end{pmatrix}$$

V našem primeru je gradientni spust:

$$\nabla \varepsilon = \begin{pmatrix} \frac{\partial \varepsilon}{\partial A} & \frac{\partial \varepsilon}{\partial B} & \frac{\partial \varepsilon}{\partial C} & \frac{\partial \varepsilon}{\partial D} & \frac{\partial \varepsilon}{\partial E} & \frac{\partial \varepsilon}{\partial F} \end{pmatrix}$$

Ker iščemo minimum funckije se bomo premikali v nasprotno smer gradienta. Spremembo spremenljivke A lahko zapišemo kot:

$$\Delta A = -(\nabla \varepsilon)_1 \cdot \lambda$$

kjer λ predstavlja velikost permika.

Za ostale spremenljivke velja podobno:

$$\Delta B = -(\nabla \varepsilon)_2 \cdot \lambda$$

$$\Delta C = -(\nabla \varepsilon)_3 \cdot \lambda$$

$$\Delta D = -(\nabla\varepsilon)_4 \cdot \lambda$$
$$\Delta E = -(\nabla\varepsilon)_5 \cdot \lambda$$
$$\Delta F = -(\nabla\varepsilon)_6 \cdot \lambda$$

Naše nove spremenljivke so:

$$A = A + \Delta A$$

$$B = B + \Delta B$$
$$C = C + \Delta C$$
$$D = D + \Delta D$$
$$E = E + \Delta E$$
$$F = F + \Delta F$$

S tem smo se premaknili za en korak bližje minimumu funkcije, to pomeni enačbi elipse, ki jo iščemo. Ta postopek ponovimo še n-krat. Večkrat ko ga ponovimo, bolj natančen bo naš rezultat.

Natančnost našega rezultata je odvisna tudi od λ (velikost premika) in začetnih vrednost spremenljivk A, B, C, D, E in F. Z manjšo velikostjo premika bo rezultat bolj natančen, vendar bomo morali postopek večkrat ponoviti. Ker s postopkom iščemo samo lokalne minimume, nam začetna vrednost spremenljivk določi kateri minimum bomo našli.