Matris Türevleri

Gradyan

m boyutlu vektor x'i alan ve geriye tek sayi sonucu donduren bir $f(x)$ fonksiyonunun x'e gore turevini nasil aliriz? Yani $x \in \mathbb{R}^m$ ve bir vektor,

$$x = \left[\begin{array}{c} x_1 \\ \vdots \\ x_m \end{array}\right]$$

Bu durumda x'in her hucresine / ogesine gore kismi turevler (partial derivatives) alinir, sonucta tek boyutlu / tekil sayili fonksiyon, turev sonrasi m boyutlu bir sonuc vektorunu yaratir, yani

$$\frac{\partial f}{\partial x} = \left[\begin{array}{c} \frac{\partial f}{\partial x_1} \\ \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_m} \end{array}\right]$$

Bu sonuc tanidik gelmis olabilir, bu ifade gradyan olarak ta bilinir.

$$\frac{\partial f}{\partial x} = \nabla f = grad\ f(x)$$

Elde edilen vektor surpriz degil cunku tek, skalar bir deger veren bir fonksiyonun x icindeki *her ogensinin* nasil degistigine gore bunun fonksiyon uzerindeki etkilerini merak ediyorduk, ustteki vektor oge bazinda bize aynen bunu gosteriyor. Yani tek skalar sonuc m tane turev sonucuna ayriliyor, cunku tek sonucun m tane secenege gore degisimini gormek istedik. Not olarak belirtelim, gradyan vektoru matematiksel bir rahatliktir, bir kisayoldur, bir ziplama noktasidir, yani matematiksel olarak turetilerek ulasilan ana kurallardan biri denemez. Fakat cok ise yaradigina suphe yok.

Tek Parametreye Gore Matris Turevi

Eger bir A matrisinin tum ogeleri bir θ parametresine bagli ise, o matrisin θ'ya gore turevi icin tum elemanlarinin teker teker θ'ya gore turevleri alinir,

$$\frac{\partial A}{\partial \theta} = \left[\begin{array}{cccc} \frac{\partial a_{11}}{\partial \theta} & \frac{\partial a_{12}}{\partial \theta} & \dots & \frac{\partial a_{1n}}{\partial \theta} \\ \frac{\partial a_{21}}{\partial \theta} & \frac{\partial a_{22}}{\partial \theta} & \dots & \frac{\partial a_{2n}}{\partial \theta} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial a_{m1}}{\partial \theta} & \frac{\partial a_{m2}}{\partial \theta} & \dots & \frac{\partial a_{mn}}{\partial \theta} \end{array}\right]$$

Cok Parametreli Matris Turevi

Simdi ilginc bir varyasyon; diyelim ki hem fonksiyon $f(x)$'e verilen x cok boyutlu, hem de fonksiyonun sonucu cok boyutlu! Bu gayet mumkun bir durum. Bu durumda ne olurdu?

Eger kismi turevlerin her turlu oge degisimini temsil etmesini istiyorsak, o zaman hem her girdi hucresi, hem de her cikti hucresi icin bu degisimi saptamaliyiz. Jacobian matrisleri tam da bunu yapar. Eger m boyutlu girdi ve n boyutlu cikti tanimlayan f'in turevini almak istersek,

$$
J(x) = \frac{\partial f(x)}{\partial x} = \left[\begin{array}{ccc} \frac{\partial f_1(x)}{\partial \theta} & \dots & \frac{\partial f_1(x)}{\partial \theta} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(x)}{\partial \theta} & \dots & \frac{\partial f_n(x)}{\partial \theta} \end{array} \right]
$$

Vektor Turevleri

Eger bir $x \in \mathbb{R}^m$ vektorunden bagimsiz bir A matrisi o x ile carpiliyor ise, bunlarin x'e gore turevi nedir?

$$
\frac{\partial}{\partial x^T} [Ax] = A
$$

Ustteki sonuc aslinda tek sayili / boyutlu ortamda $2x$ gibi bir ifadenin x'e gore turevini alinca 2 elde etmeye esdeger. Ispat icin soyle dusunelim, eger $a_i \in \mathbb{R}^n$ ise (ki devrigi alininca bu vektor yatay hale gelir, yani altta bu yatay vektorleri ust uste istifliyoruz),

$$
A = \left[\begin{array}{c} a_1^T \\ \vdots \\ a_m^T \end{array} \right]
$$

Bu durumda Ax ne olur? *Matris Carpimi* yazisindaki satir bakis acisi dusunulurse, A'in bir satirinin her ogesi x'in tum satirlarini (burada x vektor oldugu icin her satir tek bir sayidan ibaret) kombine ederek o sonuc satirini olusturmaktadir, o zaman

$$
A = \left[\begin{array}{c} a_1^T x_1 \\ \vdots \\ a_m^T x_m \end{array} \right]
$$

Kaynaklar

Duda, Hart, *Pattern Classification*