Շարժվող միջին

Հայկ Կարապետյան

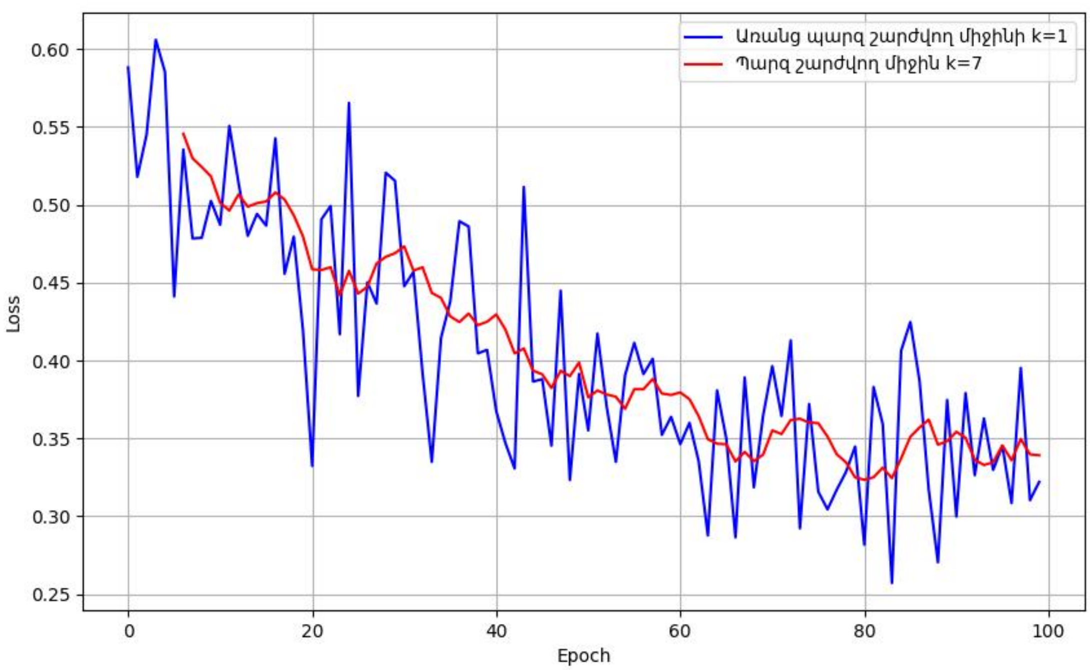
Շարժվող միջինը վիճակագրական հաշվարկ է, որն օգտագործվում է տվյալների կետերը վերլուծելու համար՝ բոլոր տվյալների կետերից ստեղծելով տարբեր միջինների շարքեր։ Այն սովորաբար օգտագործվում է կարճաժամկետ տատանումները հարթելու համար։ Դիտարկենք շարժվող միջինի երեք տեսակ։

1 Պարզ շարժվող միջին

Տրված կետում պարզ շարժվող միջինը (simple moving average), դա նախորդ k հատ հաջորդական կետերի թվաբանական միջինն է։ Եթե ունենք $x_1, x_2, \dots x_n$ հաջորդական կետեր, ապա պարզ շարժվող միջինը հաշվելուց հետո կստանանք հետևյալ կետերը.

$$\mu_n = \frac{x_{n-k+1} + \cdots + x_n}{k}, n = k, k+1, \dots 9$$

k-ն ցույց է տալիս, թե նախորդ կետերից քանի հատն է պետք հաշվի առնել նոր կետի ստացման ժամանակ։ Այն որոշում ենք մենք։ Օրինակ՝ կորստի ֆունկցիան շատ է տատանվում և մենք ցանկանում ենք ավելի քիչ տատանվող գրաֆիկ ստանալ։ Կարող ենք վերցնել ամեն կետի արժեքը՝ սկսած 7-րդ կետից, և փոխարինել նախորդ 7 կետերի միջին արժեքով (Նկար 1)։ Ինչքան մեծ վերցնենք k-ի արժեքը, այնքան ավելի սահուն կստացվի գրաֆիկը։



Նկար 1: Կորստի ֆունկցիայի գրաֆիկը պարզ շարժվող միջին կիրառելուց առաջ և հետո

2 Կուտակային շարժվող միջին

n-րդ կետում կուտակային շարժվող միջինը (cumulative moving average), դա նախորդ n-1 հատ հաջորդական կետերի թվաբանական միջինն է: Եթե ունենք $x_1, x_2, \dots x_n$ հաջորդական կետեր, ապա կուտակային շարժվող միջինը հաշվելուց հետո կստանանք հետևյալ կետերը.

$$\mu_n = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Ստանանք ռեկուրենտ բանաձև՝ ձևափոխելով առկան.

$$\mu_n = \frac{(x_1 + x_2 + \dots + x_{n-1}) + x_n}{n} = \frac{(n-1)\,\mu_{n-1} + x_n}{n} =$$
$$= \left(1 - \frac{1}{n}\right)\mu_{n-1} + \frac{1}{n}x_n$$

Սկզբնական բանաձևում երևում է, որ բոլոր կետերը (x-երը) հաշվի ենք առնում նույն $\frac{1}{n}$ կշռով: Ռեկուրենտ բանաձևը ցույց է տալիս, որ n-րդ պահին $x_1, x_2, \dots x_{n-1}$ կետերին ավելի մեծ ուշադրություն ենք դարձնում, քան x_n կետին: n-ի մեծ արժեքների դեպքում $\left(1 - \frac{1}{n}\right)$ արժեքը ձգտում է 1-ի իսկ, $\frac{1}{n}$-րդի արժեքը ձգտում է 0-ի: Ի տարբերություն նախորդի, այս դեպքում հաշվի են առնվում բոլոր կետերը, բայց եթե մեզ համար նոր եկած կետերն կարևոր են և պետք չէ դրանց շատ մոտիկացնել 0-ի (օրինակ՝ դոլարի գնի կանխատեսում), ապա կուտակային շարժվող միջինը լավ տարբերակ չէ:

3 Էքսպոնենցիալ շարժվող միջին

Եթե ունենք հետևյալ հաջորդական կետերը $x_1, x_2, \dots x_n$, ապա էքսպոնենցիալ շարժվող միջինը (exponential moving average) հաշվելուց հետո կստանանք հետևյալ կետերը.

$$\mu_n = \alpha\mu_{n-1} + (1-\alpha)\,x_n \quad n \geq 2,\ 0 < \alpha < 1$$

Հետևյալ ռեկուրենտ բանաձևը պարզեցնենք և տեսնենք, թե ամեն կետը ինչ գործակցով է մասնակցում:

$$\mu_n = \alpha\mu_{n-1} + (1-\alpha)x_n = \alpha(\alpha\mu_{n-2} + (1-\alpha)x_{x-1}) + (1-\alpha)x_n =$$
$$= \alpha^2\mu_{n-2} + (1-\alpha)\alpha x_{n-1} + (1-\alpha)x_n =$$
$$= \alpha^3\mu_{n-3} + (1-\alpha)\alpha^2 x_{n-2} + (1-\alpha)\alpha x_{n-1} + (1-\alpha)x_n =$$
$$= \alpha^{n-1}\mu_1 + (1-\alpha)\alpha^{n-2}x_2 + (1-\alpha)\alpha x_{n-1} + (1-\alpha)x_n$$

Կարող ենք տեսնել, որ կետը ինչքան հին է, այնքան գործակիցն ավելի փոքր է, այսինքն հիշողության մեջ ավելի քիչ ենք հաշվի առնում: Ամեն նոր տվյալ, հաշվի ենք առնում նույն չափով՝ $(\alpha-1)$ գործակցով, անկախ հիշողության քանակից: Շարժվող միջինը կոչվում է էքսպոնենցիալ, քանի որ ամեն նախորդ կետի գործակիցը բազմապատկվում է α-ով և էքսպոնենցիալ կերպով փոքրանում է: Նկատենք որ α-ի աստիճանի և x-ի ինդեքսի գումարը միշտ հավասար է n: Երբ n-ը մեծանում է՝ $\alpha^{n-1}x_1$-ը ձգտում է 0-ի և գրեթե ազդեցություն չի ունենում նոր եկած տվյալի վրա: α-ի արժեքը վերցնում ենք 0.5-ից ավելի մեծ, որպեսզի հիշողությանը ավելի շատ ուշադրություն դարձնենք, քան նոր եկած տվյալին, քանի որ նոր եկած տվյալը կարող է լինել շեղում (outlier), իսկ հիշողության մեջ լինեն հարյուրավոր տվյալներ: