

ADATSZERKEZETEK ÉS ALGORITMUSOK

Algoritmusok hatékonysága
„Elméleti alapok”

Hatékonyság

- Milyen hatékony egy algoritmus?
 - Legtöbbször csak a lépésszám nagyságrendje érdekes.
 - Hogyan függ a lépésszám az input méretétől?
 - Az input méretét legtöbbször n -nel jelöljük.
 - A lépésszám ennek egy f függvénye, azaz ha n méretű az input, akkor az algoritmus $f(n)$ lépést végez.
 - Igazából az f függvény az érdekes.
 - $100n$ vagy $101n$, általában mindegy
 - n^2 vagy n^3 már sokszor nagy különbség, de néha mindegy
 - n^2 vagy $2n$ már mindig nagy különbség

Függvények nagyságrendje

- Definíció – A g **aszimptotikus felső korlátja** f -nek
 - Ha $f(x)$ és $g(x)$ az R^+ egy részhalmazán értelmezett valós értékeket felvevő függvények, akkor $f = \mathcal{O}(g)$ jelöli azt a tényt, hogy vannak olyan $c, k > 0$ állandók, hogy $|f(x)| \leq c * |g(x)|$ teljesül, ha $x \geq k$.
- Például:
 - $100n + 300 = \mathcal{O}(n)$
 - hiszen $k = 300$; $c = 101$ -re teljesülnek a feltételek
 - $100n + 300 \leq 101n$, ha $n \geq 300$

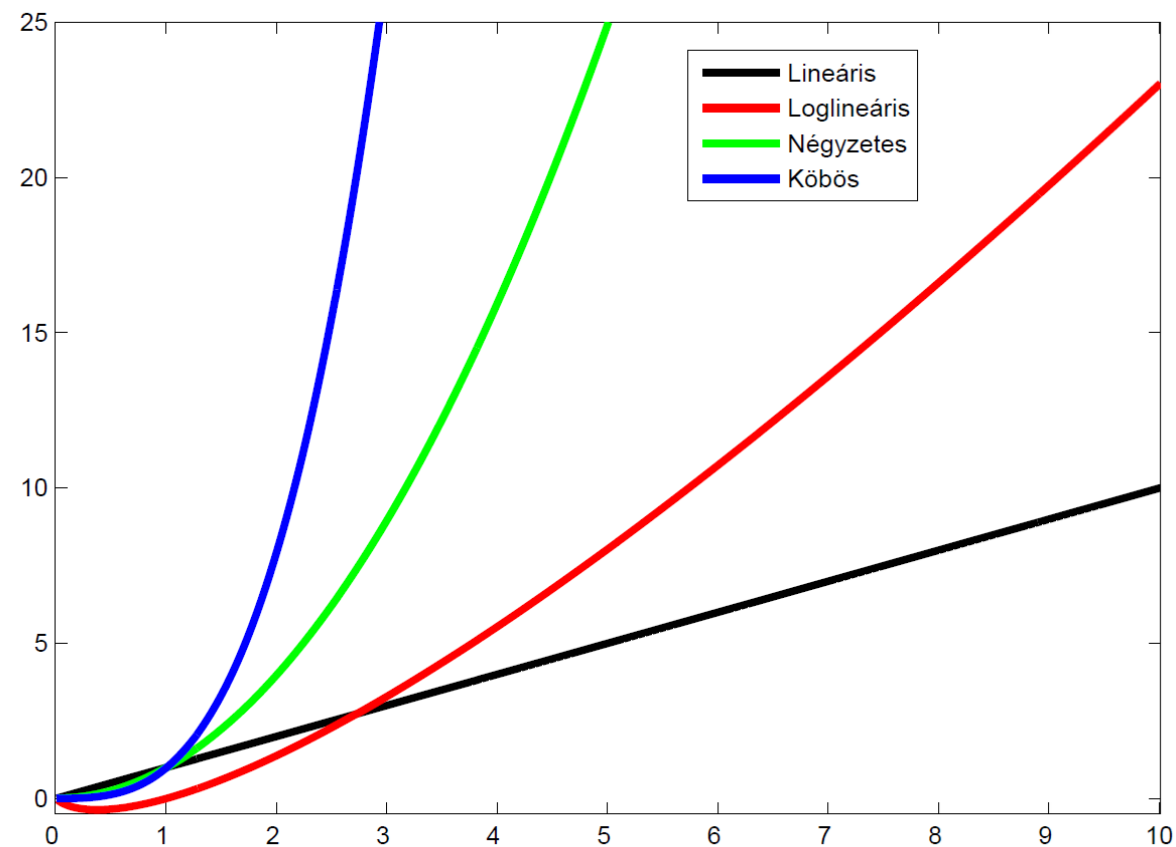
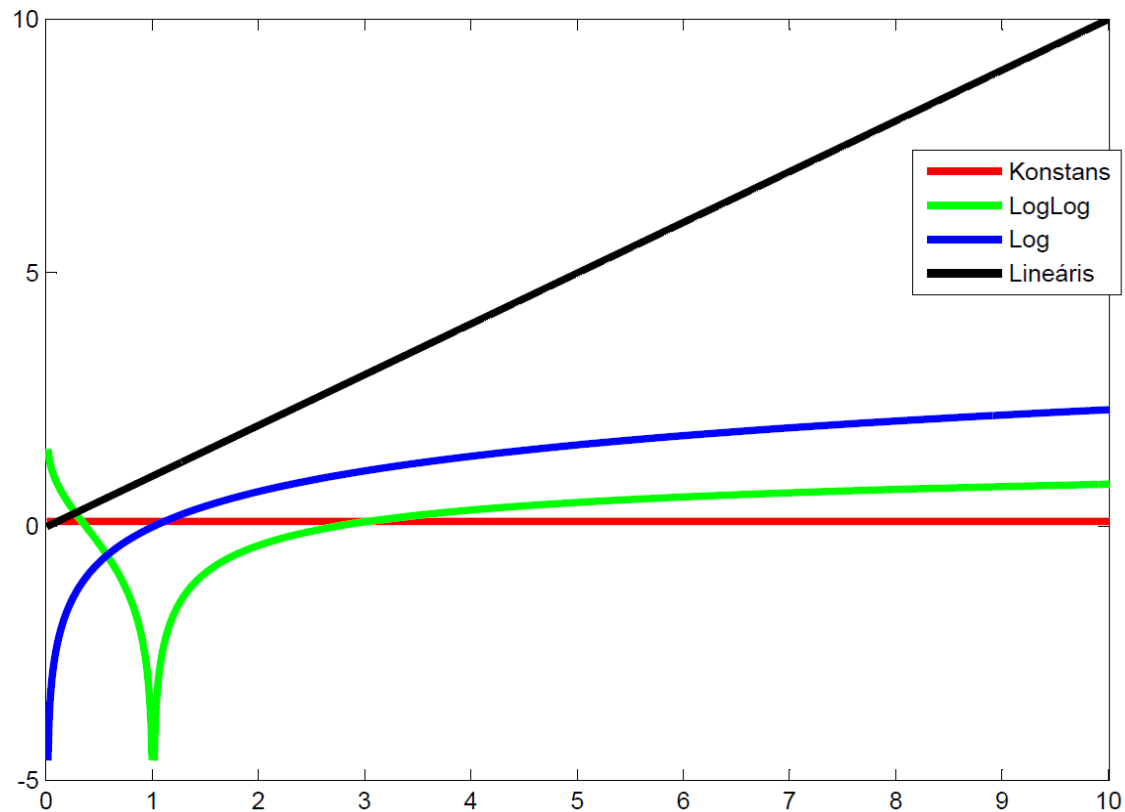
Függvények nagyságrendje

- Definíció – A g **aszimptotikus alsó korlátja** f -nek
 - Ha $f(x)$ és $g(x)$ az R^+ egy részhalmazán értelmezett valós értékeket felvevő függvények, akkor $f = \Omega(g)$ jelöli azt a tényt, hogy vannak olyan $c, k > 0$ állandók, hogy $|f(x)| \geq c * |g(x)|$ teljesül, ha $x \geq k$.
- Például:
 - $100n - 300 = \Omega(n)$
 - hiszen $k = 300$; $c = 99$ -re teljesülnek a feltételek

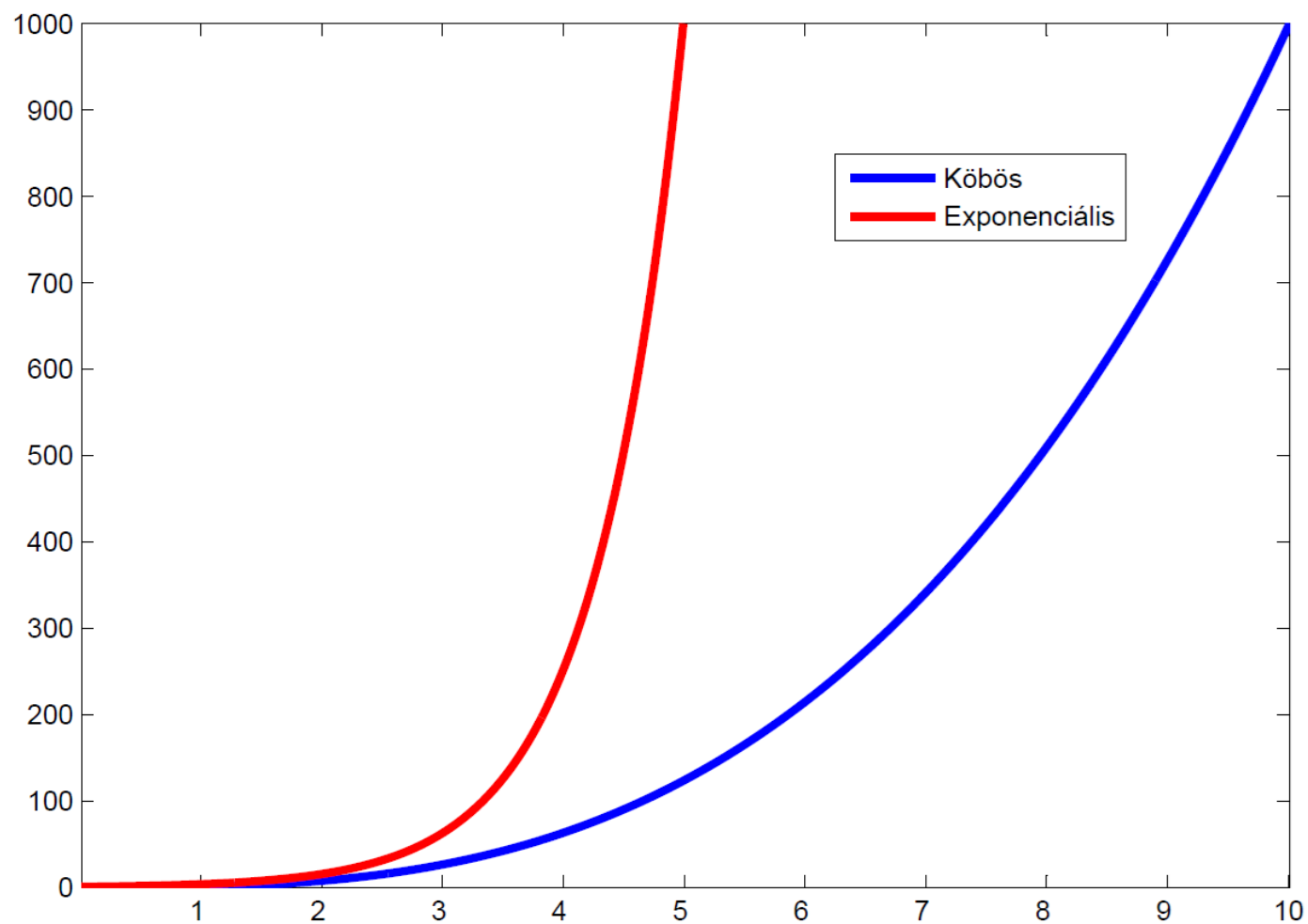
Függvények nagyságrendje

- Definíció – A g **aszimptotikus éles korlátja** f -nek
 - Ha $f = \Omega(g)$ és $f = \mathcal{O}(g)$ egyaránt teljesül, akkor $f = \Theta(g)$
- Például:
 - $100n - 300 = \Theta(n)$

Példa



Példa

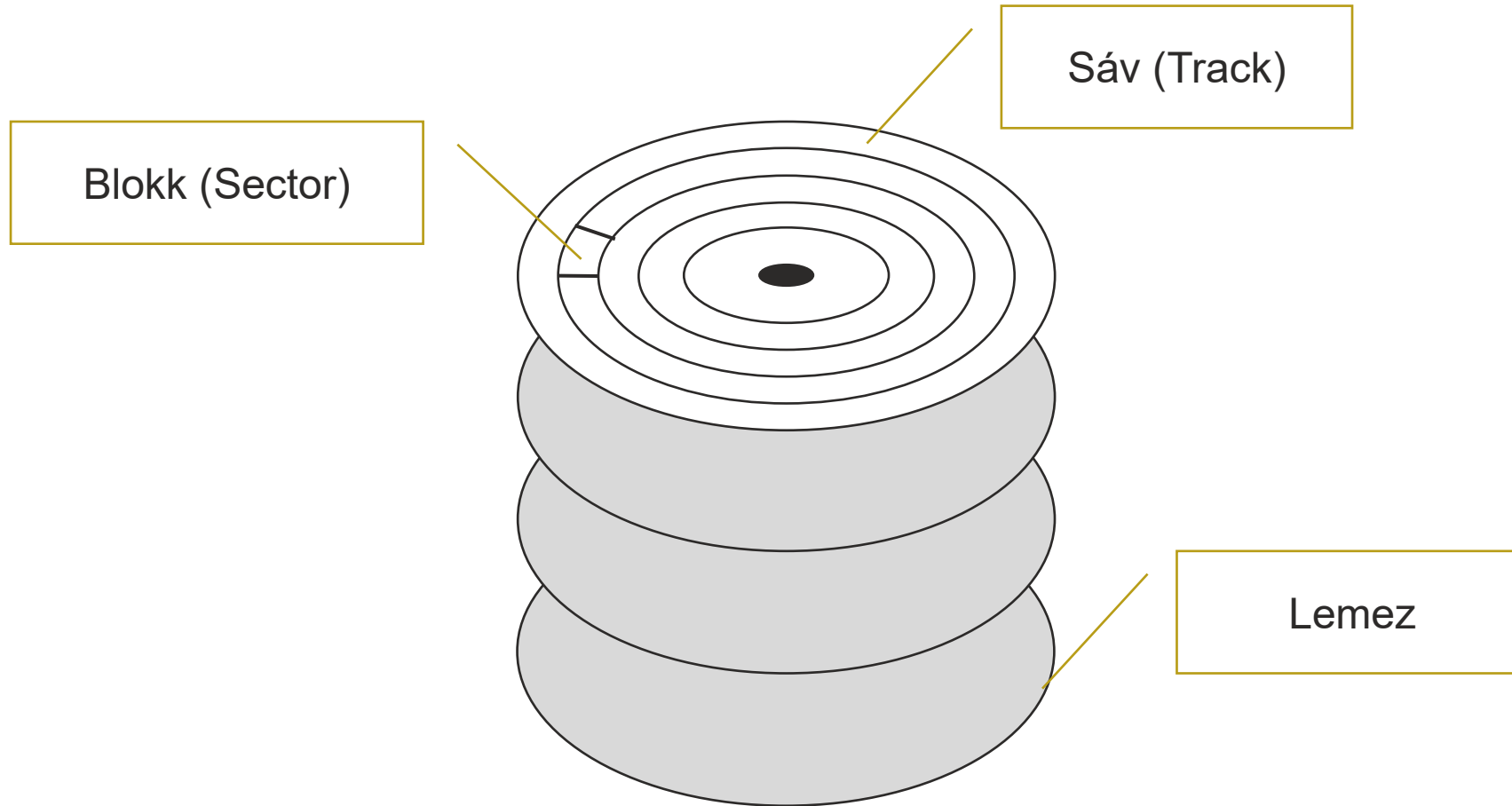


„Külső” algoritmusok

Adattárolás sémája

- Az adattárolás egy háttértárolón (lemezen) nem bit/byte egységbe van szervezve
 - Nagyobb egységek: lap (HDD: szektor, SSD: blokk, NAS: hálózati csomag)
 - Ez lehet például 2048 byte, vagy 4096 byte
 - Ez az átvitel egysége
 - Lényegében az átvitelek száma határozza meg a sebességet
 - A háttértárról egy lap olvasása lassabb, mint a főmemória (RAM) olvasása
 - ~ms vs. ~ns
 - Még SSD esetén is!

Mágnestlemez esetén



Külső algoritmusok

- Az eddig látott rendezéseknél feltettük, hogy az adatok a központi memóriában vannak
 - Ennek megfelelt, hogy a hatékonyságot az összehasonlítások számában mértük.
- Ha az adatok háttértárban vannak, akkor a futási idő döntő részét az I/O utasítások teszik ki.
 - Az I/O egysége az 1 blokk, ami $k * 512$ byte valamely kis k -val, pl. 1024 v. 2048 byte.
 - A hatékonyságot a szükséges **blokk I/O**-k számában mérjük.
- Külső rendezésre igazából csak az összefésüléses rendezés (MergeSort) alkalmas.
 - Miért?