Približno rješavanje sustava jednadžba (skica)

1. Uvod

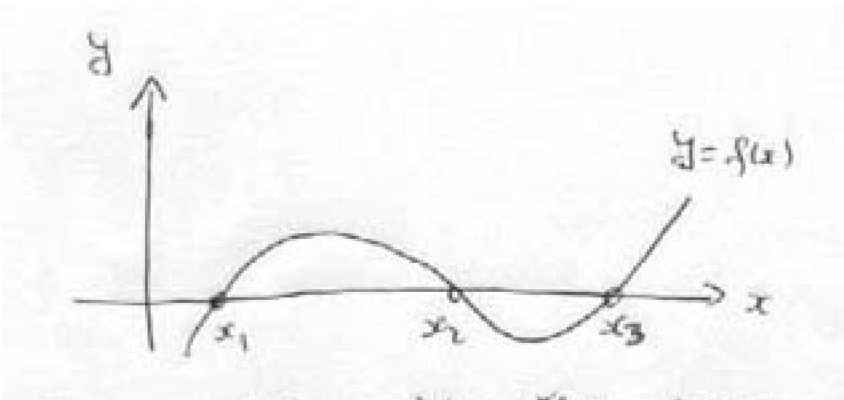
Većina se inženjerskih problema svodi na rješavanje jednažba: s jednom ili više nepoznanica, sustava jednadžba, matričnih jednadžba, diferencijalnih jednadžba, integralnih jednadžba itd. U pravilu, takve se jednadžbe ne mogu egzaktno riješiti. Zato se postavlja problem približnog rješavanja jednadžba. Ovdje ćemo skicirati Newtonovu metodu (Ralphson-Newtonovu metodu) za približno rješavanje jednadžba s jednom nepoznanicom i sustava jednadžba.

Jednadžba s jednom nepoznanicom je jednadžba oblika:
$f(x)=0$,
gdje je *f* funkcija jedne varijable. Tu se ograničavamo na slučaj kad je *f* realna funkcija realne varijable definirana na nekom intervalu realnih brojeva, koja je **neprekinuta** na tom intervalu (tj. graf joj je neprekinuta crta), štoviše, u pravilu mislimo na **elementarne funkcije** *f* (tj. na funkcije koje se mogu zadati s konačno mnogo operacija na kalkulatoru), posebice te funkcije imaju derivacije bilo kojeg reda.

Rješenje jednadžbe je bilo koji realni broj x^* takav da je $f(x^*)=0$.
Može se dogoditi da jednadžba nema rješenja, da ima konačno mnogo rješenja, ili da ima beskonačno mnogo rješenja.

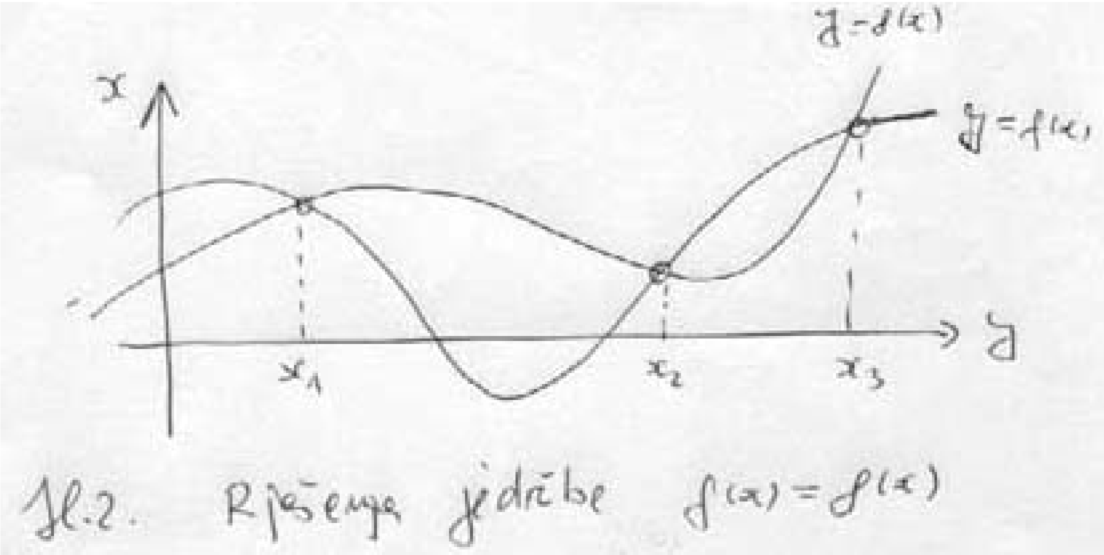
Primjer 1. (i) Jednadžba $x^2+1=0$ nema (realnih) rješenja.
(ii) Jednadžba $x^2-1=0$ ima dva rješenja: $x_1 = -1$ i $x_2 = 1$.
(iii) Jednadžba $\sin x = 0$ ima beskonačno mnogo rješenja: realne brojeve $k\pi$, gdje k prolazi skupom cijelih brojeva.

Grafička interpretacija rješenja. Realna rješenja jednadžbe $f(x)=0$ jesu upravo oni realni brojevi u kojima graf funkcije *f* siječe x-os (sl.1).

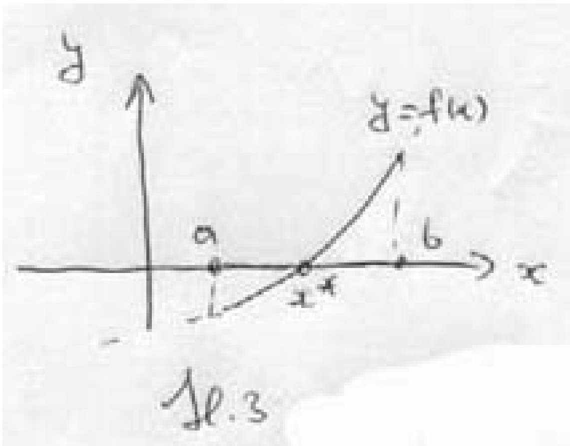


Sl.1. Rješenja jednadžbe $f(x)=0$

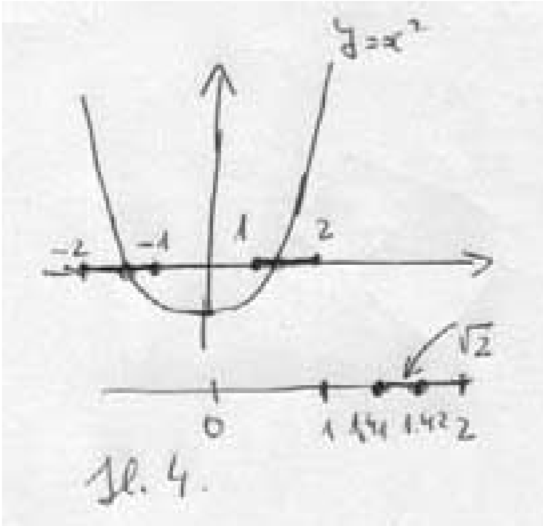
Varijanta: Realna rješenja jednadžbe $f(x)=g(x)$ jest skup prvih koordionata točaka u kojima se sijeku grafovi funkcija *f* i *g* (sl.2.).



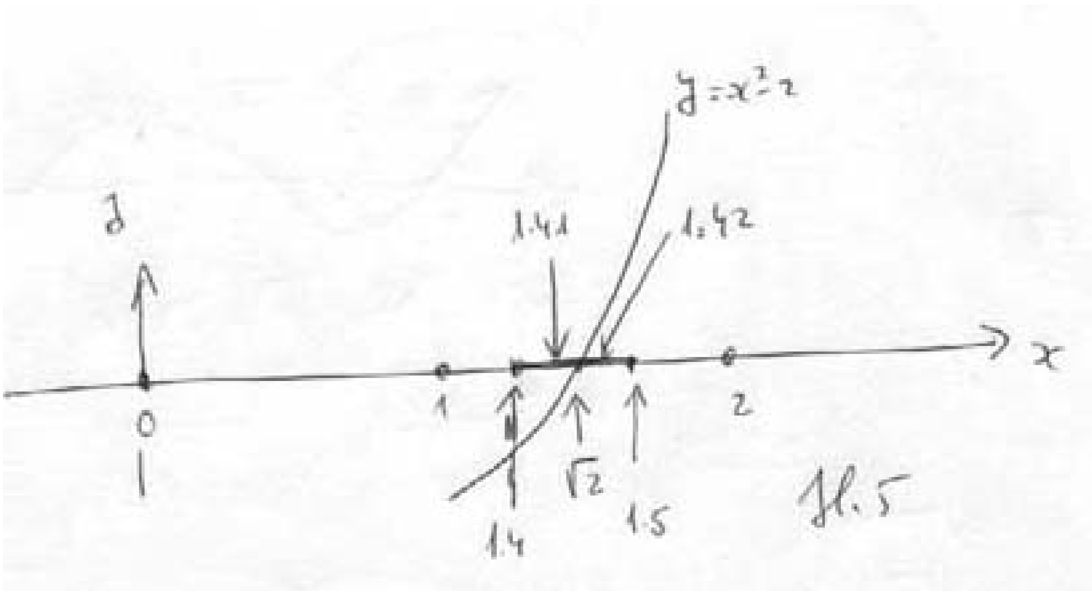
Interval izoliranosti rješenja. To je svaki interval [a,b] unutar kojega ima točno jedno rješenje jednadžbe (sl.3.).



Primjer 2. Interval [-2,-1] je interval izoliranosti rješenja $-\sqrt{2}$ jednadžbe x^2-2=0, a interval [1,2] je interval izoliranosti rješenja $\sqrt{2}$ te iste jednadžbe (sl.4.). Naravno da ima i užih intervala izoliranosti od tih. Na primjer, [1.41, 1.42] takodjer je interval izoliranosti rješenja $\sqrt{2}$ jer je $1.41 < \sqrt{2} < 1.42$.



Približno rješenje jednadžbe f(x)=0. To je svaki broj α za koji vrijedi $f(\alpha) \approx 0$.
Na primjer, 1.4 je približno rješenje jednadžbe $x^2-2=0$. Naime,
$1.4^2-2= 1.96-2 = - 0.04 \approx 0$.
Kažemo da je 1.4 **aproksimacija** rješenja $\sqrt{2}$ te jednadžbe jer je $\sqrt{2} \approx 1.41$.
Takodjer, 1.41 je približno rješenje te jednadžbe, jer je
$1.41^2-2= 1.9881-2 = - 0.0119 \approx 0$.
Vidimo da je 1.41 **bolja aproksimacija** te jednadžbe od aproksimacije 1.4 ; to se očituje time što je 1.41 bliže broju $\sqrt{2}$ nego što je 1.4 (sl.5.).

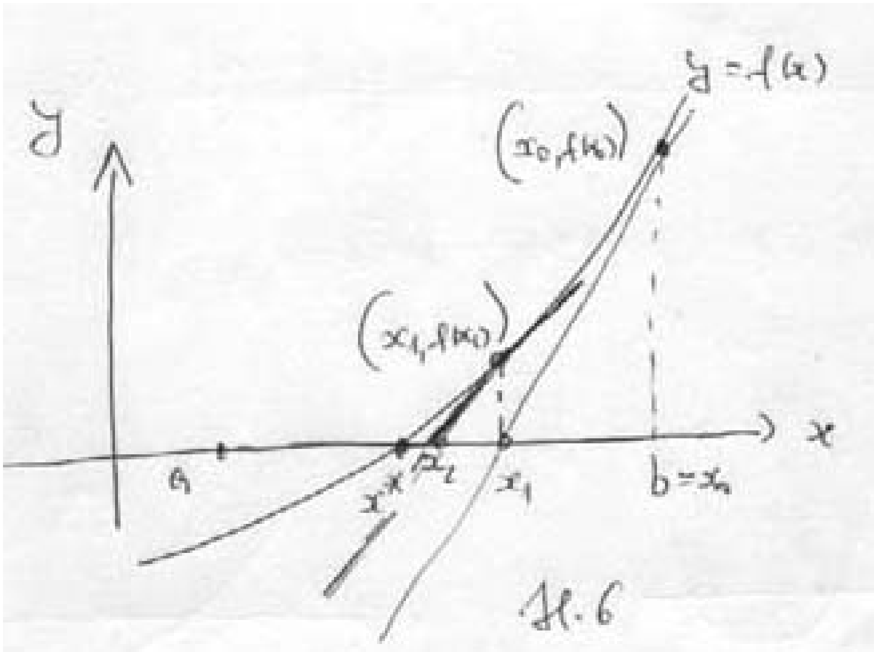


Pogrješka aproksimacije. To je udaljenost ε stvarnog rješenja x^* i približnog rješenja (aproksimacije) α, dakle:
$\varepsilon := |x^* - \alpha|$.
U pravilu ne znamo točno pogrješku aproksimacije, jer bismo onda znali i rješenje. Zato je potrebno znati neku ocjenu za pogrješku. Ta se ocjena zove **ocjena grješke**. Na primjer, za aproksimaciju 1.4 od $\sqrt{2}$, vrijedi da je $\varepsilon < 0.1$ (to je ocjena grješke), jer je
$1.4 < \sqrt{2} < 1.5$.

Metoda približnog rješavanja jednadžba. To je svaka metoda kojom se startajući od **nulte aproksimacije** x_0, unutar intervala izoliranosti rješenja x^*, redom dobivaju sve bolje aproksimacije $x_1, x_2, \ldots$ koje se približavaju (konvergiraju) pravom rješenju x^*.

2. Metoda tangente (Newtonova metoda) za približno rješavanje jednadžbe s jednom nepoznanicom.

Ta se metoda geometrijski zasniva na povlačenju tangente na graf funkcije *f* koja sudjeluje u jednadžbi $f(x)=0$ (sl.6.).



Analitički, nakon biranja nulte aproksimacije (x_0=a ili x_0=b, za interval izoliranosti [a,b]), dobijemo:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} \quad \text{i, općenito,}$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Jedan pogled na formulu za metodu tangente. Gornju formulu možemo izvesti tako da anlitički odredimo presjek tangente na graf funkcije f u točki $(x_0,f(x_0))$ s osi x. Tako dobijemo x_1 u ovisnosti o x_0, a onda analogno napišemo opću formulu. Taj pogled nije previše pogodan za popćenje na slučaj sustava jednadžba (iako je i to moguće). Zgodnije je razmišljati da smo, funkciju f zamijenili njenom linearnom aproksimacijom oko x_0, tj.
$f(x) \approx f(x_0)+f'(x_0)(x-x_0)$ za $x \approx x_0$,
a potom umjesto (u pravilu komplicirane) jednadžbe $f(x)=0$, gledati linearnu jednadžbu

$f(x_0)+f'(x_0)(x-x_0)=0$,
u kojoj smo f zamijenili njenom linearnom aproksimacijom. Linearnu jednadžbu klako je riješiti; označimo njeno rješenje s x_1, dakle

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

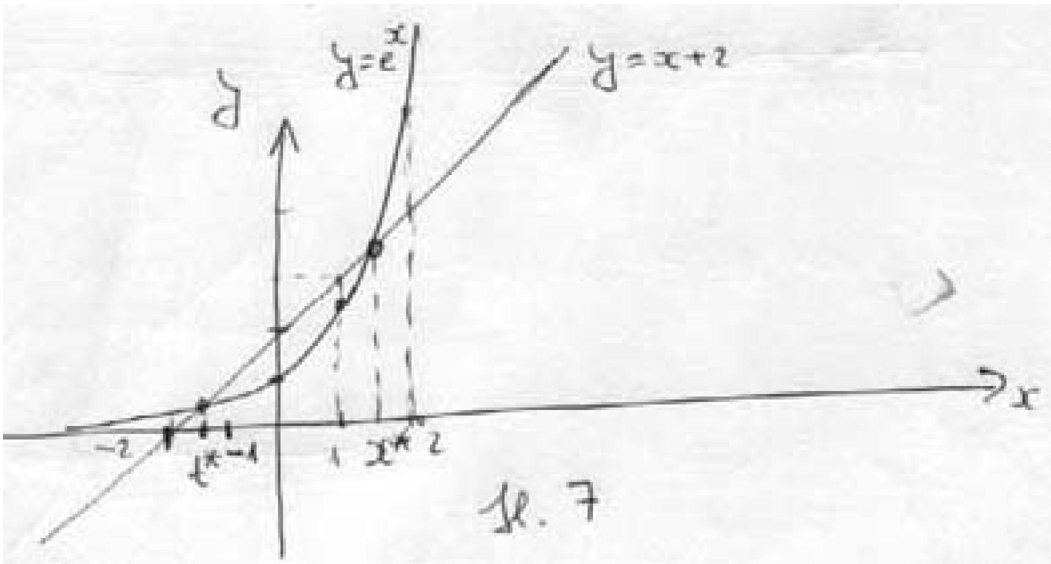
kako smo već napisali. Naravno x_1 nije ujedno i rješenje jednadžbe f(x)=0, već samo približno (**prva aproksimacija**), a ako pogodno izaberemo nultu, onda će prva aproksimacija biti bolja od druge, treća bolja od druge itd.

Primjer 3. Metodom tangente približno riješimo jednadžbu $e^x = x + 2$.

1. korak. Izolacija rješenja.
Provodimo je **grafičkom** i **analitičkom metodom**.
Grafička metoda: nacrtamo u istom koordinatnom sustavu krivulje s jednadžbama:
$y=e^x$ i $y = x+2$.
Vidimo da se te krivulje sijeku u dvjema točkama, zato jednadžba ima dva rješenja (sl.7.). Prvo je rješenje između –2 i –1, a drugo između 1 i 2.

Analitička metoda

x	e^x-x-2
0	-1
1	-0.28
2	3.39
-1	-0.6
-2	0.13



S objema metodama dolazimo do istog zaključka o broju rješenja.

2. korak. Računanje 1. i 2. derivacije funkcije $f(x) := e^x$-x-2.
$f'(x) = e^x-1$; $f''(x) = e^x$.

3. korak. Provjeravanje monotonosti funkcija f', f'' na dobivenim intervalima, potom njihovim neponištavanjem na tim intervalima.
Polazimo od druge derivacije (jer pomoću nje lakše možemo ocijeniti prvu; iako u ovom primjeru to nije potrebno jer sve teče glatko).
Očito je f'' rastuća i pozitivna svugdje, posebice na tim intervalima.
f' je također rastuća (što vidimo izravno, ali i iz toga što je f'' svugdje pozitivna);

Dalje: $f'(x) > 0$, za sve x iz [1,2] (što je zbog rasta dovoljno provjeriti u x=1)
$f'(x) < 0$, za sve x iz [-2,-1] (što je zbog rasta dovoljo provjeriti u x= -1).

4. korak. Određivanje minimuma apsolutnih vrijednosti prve derivacije te maksimuma apsolutnih vrijednosti 2. derivacije na tim intervalima. Točnije:
$m<\min|f'(x)|$, za x u pojedinom intervalu
$M>\max|f''(x)|$, za x u pojedinom intervalu.
(ove smo znakove izabrali jer će nam se *m* nalaziti u nazivniku određene formule, a *M* u brojniku te formule).
Moramo posebno raditi na svakom od intervala. Iz koraka 3. znademo da su f' i f'' monotone na tim intervalima i da na njima ne mijenjaju predznak. To nam omogućuje da m,M izaberemo računajući samo vrijednosti funkcija u rubovima (inače tako nebismo smjeli).
Interval [1,2]. Skicirajmo grafove od f' i f''.
Zaključujemo:
$\min|f'(x)| = e-1$, pa možemo uzeti $m=1.7$
(približni smo rezultat dobili
zaokruživanjem na niže, odnosno odbacivanjem znamenaka).
$\max|f''(x)|=e^2$, pa možemo uzeti $M=7.4$
(zaokruživanje na više).

Interval [-2,-1]. Opet skicirajmo grafove.
Zaključujemo:
$\min|f'(x)| = 1-e^{-1}$, pa možemo uzeti $m=0.6$

$\max|f''(x)|=e^{-1}$, pa možemo uzeti $M= 0.4$.

5. korak. Određivanje uvjeta za zadanu točnost.
Iz formule $|x^*-x_{n+1}|<\frac{f^2(x_n)}{2}\cdot\frac{M}{m^3}$
i zahtjeva u zadatku: $|x^*-x_{n+1}|<0.001$,
vidimo da je dovoljno uzeti da bude: $\frac{f^2(x_n)}{2}\cdot\frac{M}{m^3}<0.001$, odnosno

$$f^2(x_n)<\frac{2m^3}{M}\cdot 0.001$$

Taj uvjet treba posebno računati za svaki od intervala.

Interval [1,2].
Dobijemo:
$f^2(x_n)<0.003$ (uzimamo manji broj, odnosno, zaokružujemo na niže).

Interval [-2,-1]
Dobijemo:
$f^2(x_n)< 0.0037$

Smisao ove ocjene je sljedeći: kad dođemo da takvog n za koji će to vrijediti, onda stajemo, a za dobru približnu vrijednost uzimamo x_{n+1}.

6. korak. Biranje nulte aproksimacije x_0.
Biramo je tako da bude rubna točka intervala uz uvjet $f(x_0)f''(x_0)>0$.
Treba posebno birati za svaki od intervala.

Interval [1,2].
Dobijemo
$x_0=2$

Interval [-2,-1]
Dobijemo
$x_0=-2$

7. korak. Računanje aproksimacija pomoću formule $x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}$,

tj. pomoću formule: $$x_{n+1} = x_n - \frac{e^{x_n} - x_n - 2}{e^{x_n} - 1}$$

I to se provodi posebno za svaki interval.

Interval [1,2].

$x_1 = 2 - \dfrac{e^2 - 2 - 2}{e^2 - 1} = 1.46955$

$x_2 = 1.46955 - \dfrac{e^{1.46955} - 1.46955 - 2}{e^{1.46955} - 1} = 1.2073$

$x_3 = 1.1488$

$x_4 = 1.1462$

(Rješenje zadovoljava tražene uvjete, jer je $(e^{1.1488} - 1.1488 - 2)^2 < 0.000032 < 0.003$).

Interval [-2,-1].

$x_1 = -2 - \dfrac{e^{-2} - (-2) - 2}{e^{-2} - 1} = -1.84348$

$x_2 = -1.8414$

(Rješenje zadovoljava tražene uvjete.)

3. Sustav dviju jednadžba s dvjema nepoznanicama

Umjesto rješavanja jednadžbe s jednom nepoznanicom u praksi često dolaze sustavi više jednadžba s više nepoznanica. Na primjer:

f(x,y)=0 (1)
g(x,y)=0,

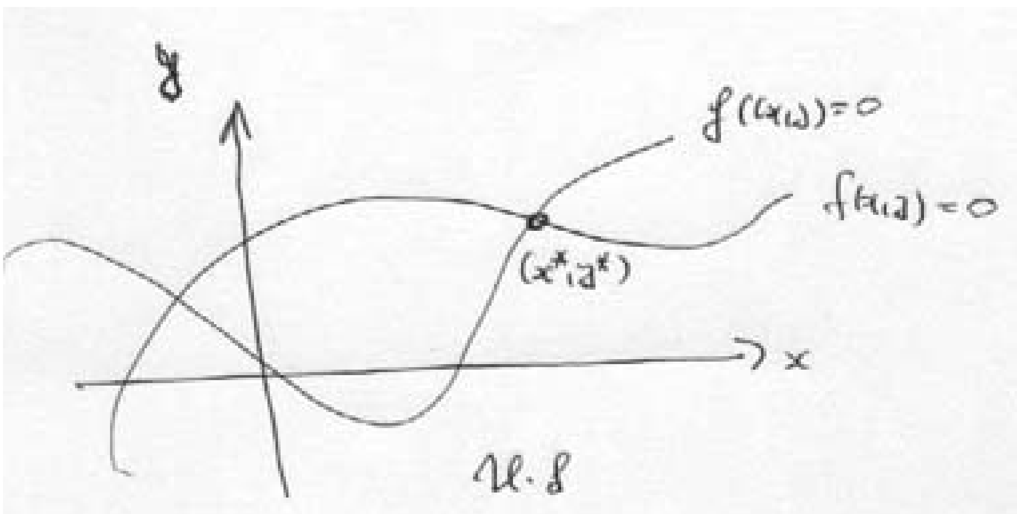
gdje su f,g funkcije dviju nezavisnih varijabla x,y je sustav dviju jednadžba s dvjema nepoznanicama. Ako je $f(x,y) := x^2 + y^2 - 10$ i $g(x,y) := 2x - y + 1$, onda sustav postaje:

$x^2 + y^2 - 10 = 0$ (2)
$2x - y + 1 = 0$.

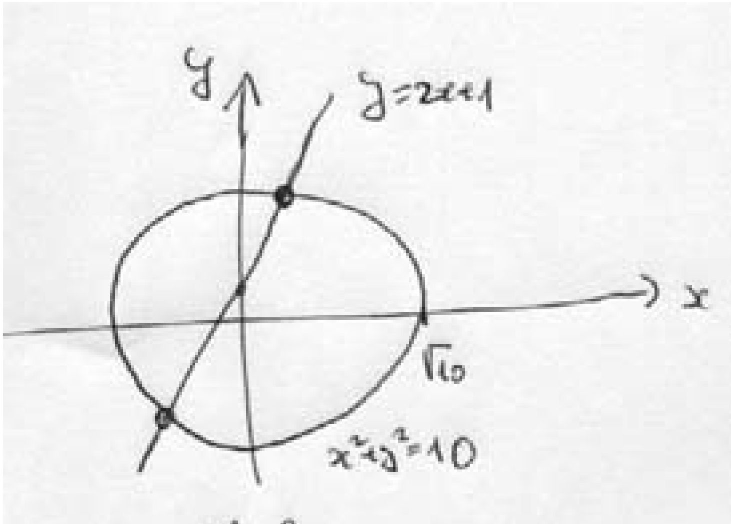
Rješenje sustava (1) je svaki uređeni par realnih brojeva (x*,y*) za koji vrijedi f (x*,y*)=0 i g(x*,y*)=0. Na primjer:

(1,3) **je rješenje** sustava (2), jer je $1^2 + 3^2 - 10 = 0$ i $2\cdot1 - 3 + 1 = 0$.
(3,1) **nije rješenje** sustava (2), jer je $3^2 + 1^2 - 10 = 0$, ali $2\cdot3 - 1 + 1 = 6 \neq 0$.

Geometrijska interpretacija rješenja sustava (1): to je presjek odgovarajućih krivulja zadanih jednadžbama f(x,y)=0, odnosno g(x,y)=0 (sl.8.).



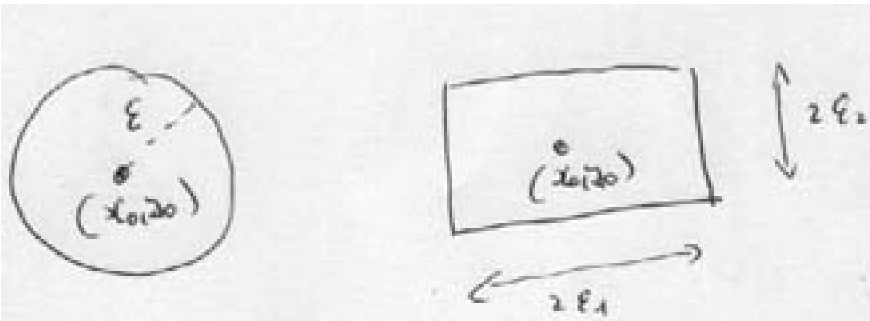
Na primjer, za sustav (2), to je presjek kružnice i pravca (sl. 9). Vidimo da sustav ima dva rješenja.



sl. 9

Približno rješenje sustava (1): to je svaki uređeni par realnih brojeva (x_0,y_0) za koji vrijedi $f(x_0,y_0) \approx 0$ i $g(x^*,y^*) \approx 0$. Na primjer:
(-2,-2.5) **je približno rješenje** sustava (2), jer je $(-2)^2+(-2.5)^2-10 = 0.25 \approx 0$ i $2\cdot(-2)-(-2.5)+1 = -0.5 \approx 0$.

Područje izoliranosti rješenja (x^*,y^*) **sustava** (1): to je svako područje u ravnini unutar kojega se nalazi (x^*,y^*), ali ni jedno drugo rješenje tog sustava. To je obično krug ili pravokutnik (sl. 10). Svaki uređeni par (x_0,y_0) unutar područja izoliranosti smatramo približnim rješenjem i u pravilu ga možemo uzeti za nultu aproksimaciju.



sl. 10

Napomena o složenijim sustavima. Analogno je sa sustavima od više jednadžba s više nepoznanica. U praksi su najvažniji sustavi koji imaju jednako jednadžba kao i nepoznanica i mićemo samo takve razmatrati. Na primjer za **linearne** takve sustave vrijedi da u pravilu imaju točno jedno rješenje. Ako ima više nepoznanica nego jednadžba, onda u pravilu ima beskonačno mnogo rješenja i skup rješenja ima dimenziju koja je razlika izmedju broja nepoznanica i broja rješenja (svaka nova jednadžba u pravilu smanjuje stupnjeve slobode, tj. smanjuje dimenziju skupa rješenja). Ako ima više jednadžba nego nepoznanica, onda je sustav prezasićen i u pravilu nema rješenja.

4. Newtonova metoda za približno rješavanje sustava dviju jednadžba s dvjema nepoznanicama.

Postupamo analogno kao kod jednadžbe s jednom nepoznanicom.

1.korak Izolacija rješenja – biranje nulte aproksimacije (x_0,y_0) koju u matričnom obliku pišemo kao jednostupčanu matricu $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$.

2. korak – određivanje prve i daljnjih aproksimacija.

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} .- J^{-1}(x_0,y_0) \begin{bmatrix} f(x_0,y_0) \\ g(x_0,y_0) \end{bmatrix}$$

gdje je **Jakobijan J**, definiran kao $J(x_0,y_0) := \begin{bmatrix} \frac{\partial f}{\partial x}(x_0,y_0) & \frac{\partial f}{\partial y}(x_0,y_0) \\ \frac{\partial g}{\partial x}(x_0,y_0) & \frac{\partial g}{\partial y}(x_0,y_0) \end{bmatrix}$

pa vidimo analogiju s formulom za jednu varijablu: tamo smo dijelili s derivacijom, a ovdje djelujemo s inverzom od jakobijana (neka vrsta dijeljenja).

Ako uvrstimo $J^{-1}(x_0,y_0) = \dfrac{1}{\det J(x_0,y_0)} \begin{bmatrix} \frac{\partial g}{\partial y}(x_0,y_0) & -\frac{\partial f}{\partial y}(x_0,y_0) \\ -\frac{\partial g}{\partial x}(x_0,y_0) & \frac{\partial f}{\partial x}(x_0,y_0) \end{bmatrix}$,

dobit ćemo $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} - \dfrac{1}{\det J(x_0,y_0)} \begin{bmatrix} \frac{\partial g}{\partial y}(x_0,y_0) f(x_0,y_0) - \frac{\partial f}{\partial y}(x_0,y_0) g(x_0,y_0) \\ -\frac{\partial g}{\partial x}(x_0,y_0) f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0) g(x_0,y_0) \end{bmatrix}$

Slično se (n+1)-va aproksimacija dobije iz n-te.

3.korak. Ocjena pogrješke, kriterij za zaustavljanje procesa.
Tu je problem ocjene pogrješke puno složeniji od analognog problema za jednu varijablu. Za kriterij zaustavljanja procesa možemo uzeti bliskost dviju uzastopnih aproksimacija (iako to može zavarati). Na primjer, u 1-udaljenosti, stajemo ako je
$|x_{n+1} - x_n| + |y_{n+1} - y_n| < \varepsilon$,
za unaprijed odabran mali pozitivan broj ε.
Slično, u 2-udaljenosti stajemo ako je
$\sqrt{(x_{n+1} - x_n)^2 + (y_{n+1} - y_n)^2} < \varepsilon$.

Naravno, mogli bismo uzeti kriterij zaustavljanja prema vrijednosti funkcija koja sudjeluju u sustavu, na primjer, stajemo ako je:
$|f(x_n,y_n)| < \varepsilon$ i $|g(x_n,y_n)| < \varepsilon$

Primjer 4. Metodom tangente približno odredimo rješenje sustava
$x^2+y^2-10=0$
$2x-y+1=0$,
koje je u III. kvadrantu.a kriterij zaustavljanja uzmimo $|f(x_n,y_n)|<0.1$ i $|g(x_n,y_n)|<0.1$.

Tu je riječ o sustavu (2) kojega bismo mogli točno riješiti, medjutim na njemu demonstriramo metodu tangente.

1.korak Izolacija rješenja – biranje nulte aproksimacije $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$. Vidjeli smo da možemo uzeti $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}=\begin{bmatrix} -2 \\ -2.5 \end{bmatrix}$.

2. korak.
Tu je $f(x,y):=x^2+y^2-10$ i $g(x,y):=2x-y+1$, pa je $\frac{\partial f}{\partial x}=2x$, $\frac{\partial f}{\partial y}=2y$, $\frac{\partial g}{\partial x}=2$, $\frac{\partial g}{\partial y}=-1$,
$f(-2,-2.5)=0.25$; $g(-2,-2.5)=-0.5$.
Zato je $J(-2,-2.5)=\begin{bmatrix} -4 & -5 \\ 2 & -1 \end{bmatrix}$, dakle $J^{-1}(-2,-2.5)=\frac{1}{14}\begin{bmatrix} -1 & 5 \\ -2 & -4 \end{bmatrix}$.
Zato je:
$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}=\begin{bmatrix} -2 \\ -2.5 \end{bmatrix}-\frac{1}{14}\begin{bmatrix} -1\cdot 0.25+5\cdot(-0.5) \\ -2\cdot 0.25-4\cdot(-05) \end{bmatrix}=\begin{bmatrix} -2 \\ -2.5 \end{bmatrix}-\frac{1}{14}\begin{bmatrix} -2.75 \\ 1.5 \end{bmatrix}=\begin{bmatrix} -1.81 \\ -2.61 \end{bmatrix}$$
Već smo dobili dosta dobro rješenje, jer nije teško vidjeti da su stvarna rješenja sustava (1,3) i (-1.8,-2.6) .
Naravno da mi općenito ne znamo točno rješenje pa se moramo osloniti na dobivene aproksimacije. Zato računamo:
$|f(-1.81,-2.61)|=|0.082|<0.1$
$|g(-1.81,-2.61)|=|-0.01|<0.1$
Zato stajemo. Da smo uzeli kriterij zaustavljanja $|f(x_n,y_n)|<0.05$ i $|g(x_n,y_n)|<0.05$, morali bismo nastaviti s traženjem druge aproksimacije (sad bi u jakobijan trebalo uvrštavati vrijednosti iz prve aproksimacije). Dobili bismo
$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}=\begin{bmatrix} -1.8079 \\ -2.6009 \end{bmatrix}$$ (na četiri decimale)
Lako je provjeriti da je
$|f(-1.8079,-2.6009)|=|0.033|<0.5$
$|g(-1.8079,-2.6009)|=|-0.015|<0.5$, pa druga aproksimacija zadovoljava kriterij zaustavljanja.

Napomena o složenijim sustavima. Sustav triju jednadžba s trima nepoznanicama slično bi se rješavao, samo što bi sad rješenje bila uredjena trojka, jakobijan bi bio 3×3 matrica itd.

Izvod formule. Pri izvodu ćemo simulirati izvod za funkcije jedne varijable samo što ćemo koristiti linearnu aproksimaciju oko (x_0,y_0) funkcije dviju varijabla.

$$f(x,y)\approx f(x_0,y_0)+\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0)+\frac{\partial f}{\partial y}(x_0,y_0)(y-x_0)$$

g(x,y) $\approx$ g(x_0,y_0)+ $\frac{\partial g}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial g}{\partial y}(x_0, y_0)(y - x_0)$

Umjesto originalnog sustava rješavamo linearni sustav

f(x_0,y_0)+$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - x_0)$ =0

g(x_0,y_0)+$\frac{\partial g}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial g}{\partial y}(x_0, y_0)(y - x_0)$ =0

Ako rješenje doznaćimo kao $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ upravo dobijemo što smo gore napisali.