Diferenciação Numérica

Prof. Wagner H. Bonat

Universidade Federal do Paraná
Departamento de Estatística
Laboratório de Estatística e Geoinformação

Sumário

Diferenciação numérica

- Derivada dá uma medida da taxa na qual a variável y muda devido a uma mudança na variável x.
- A função a ser diferenciada pode ser dada por uma função $f(x)$, ou apenas por um conjunto de pontos (y_i, x_i).
- Quando devemos usar derivadas numéricas?
  1. $f'(x)$ é dificil de obter analiticamente.
  2. $f'(x)$ é caro para calcular computacionalmente.
  3. Quando a função é especificada apenas por um conjunto de pontos.
- Abordagens para a diferenciação numérica
  1. Aproximação por diferenças finitas;
  2. Aproximar a função por uma outra função de fácil derivação.

Aproximação da derivada por diferenças finitas

- Derivada $f'(x)$ de uma função $f(x)$ no ponto $x = a$ é definida como:

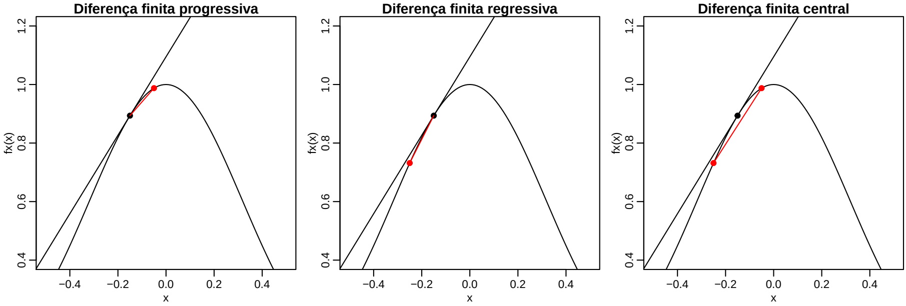
$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

- Derivada é o valor da inclinação da reta tangente à função em $x = a$.
- Escolhe-se um ponto x próximo a a e calcula-se a inclinação da reta que conecta os dois pontos.
- A precisão do cálculo aumenta quando x aproxima de a.
- Aproximação numérica: função será avaliada em diferentes pontos próximos a a para aproximar $f'(a)$.

Aproximação da derivada por diferenças finitas

- Fórmulas para diferenciação numérica:
  1. Diferença progressiva: Inclinação da reta que conecta os pontos $(x_i, f(x_i))$ e $(x_{i+1}, f(x_{i+1}))$:
  $$f'(x_i) = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}.$$
  2. Diferença regressiva: Inclinação da reta que conecta os pontos $(x_{i-1}, f(x_{i-1}))$ e $(x_i, f(x_i))$:
  $$f'(x_i) = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}.$$
  3. Diferença central: Inclinação da reta que conecta os pontos $(x_{i-1}, f(x_{i-1}))$ e $(x_{i+1}, f(x_{i+1}))$:
  $$f'(x_i) = \frac{f(x_{i+1}) - f(x_{i-1})}{x_{i+1} - x_{i-1}}.$$

Ilustração: Derivada por diferenças finitas



Aproximação da derivada por diferenças finitas

- Diferença progressiva

```
dif_prog <- function(fx, x, h) {
  df <- (fx(x + h) - fx(x))/( (x + h) - x)
  return(df)
}
```

- Diferença regressiva

```
dif_reg <- function(fx, x, h) {
  df <- (fx(x) - fx(x - h))/( x - (x - h))
  return(df)
}
```

- Diferença central

```
dif_cen <- function(fx, x, h) {
  df <- (fx(x + h) - fx(x - h))/( (x + h) - (x - h))
  return(df)}
```

Exemplo: Aproximação da derivada por diferenças finitas

- Considere $f(x) = x^3$, assim $f'(x) = 3x^2$.

```
fx <- function(x) x^3
# Diferença progressiva
dif_prog(fx, x = 2, h = 0.001)
```

```
# [1] 12.006
```

```
# Diferença regressiva
dif_reg(fx, x = 2, h = 0.001)
```

```
# [1] 11.994
```

```
# Diferença central
dif_cen(fx, x = 2, h = 0.001)
```

```
# [1] 12
```

```
# Exata
3*2^2
```

Sumário

Diferenças finitas usando expansão em série de Taylor

- As fórmulas anteriores podem ser deduzidas usando expansão em série de Taylor.
- O número de pontos para aproximar a derivada pode mudar.
- Vantagem da dedução por série de Taylor é que ela fornece uma estimativa do erro de truncamento.

Diferença finita progressiva com dois pontos

- Aproximação de Taylor para o ponto x_{i+1}

$$f(x_{i+1}) = f(x_i) + f'(x_i)h + \frac{f''(x_i)}{2!}h^2 + \frac{f'''(x_i)}{3!}h^3 + \ldots,$$

  onde $h = x_{i+1} - x_i$.
- Fixando dois termos e deixando os outros termos como um resíduo, temos

$$f(x_{i+1}) = f(x_i) + f'(x_i)h + \frac{f''(\xi)}{2!}h^2.$$

- Resolvendo para $f'(x_i)$, temos

$$f'(x_i) = \frac{f(x_{i+1}) - f(x_i)}{h} - \frac{f''(\xi)}{2!}h^2.$$

- Erro de truncamento,

$$-\frac{f''(\xi)}{2!}h^2 = O(h).$$

Diferença finita regressiva com dois pontos

- Aproximação de Taylor para o ponto x_{i-1}

$$f(x_{i-1}) = f(x_i) - f'(x_i)h + \frac{f''(x_i)}{2!}h^2 + \frac{f'''(x_i)}{3!}h^3 + \ldots,$$

  onde $h = x_i - x_{i-1}$.

- Fixando dois termos e deixando os outros termos como um resíduo, temos

$$f(x_{i-1}) = f(x_i) - f'(x_i)h + \frac{f''(\xi)}{2!}h^2.$$

- Resolvendo para $f'(x_i)$, temos

$$f'(x_i) = \frac{f(x_i) - f(x_{i-1})}{h} + \frac{f''(\xi)}{2!}h^2.$$

- Erro de truncamento,

$$\frac{f''(\xi)}{2!}h^2 = O(h).$$

Diferença finita central com dois pontos

- Aproximação de Taylor para o ponto x_{i+1}

$$f(x_{i+1}) = f(x_i) + f'(x_i)h + \frac{f''(x_i)}{2!}h^2 + \frac{f'''(\xi_1)}{3!}h^3,$$

  onde ξ_1 está entre x_i e x_{i+1}.
- Aproximação de Taylor para o ponto x_{i-1}

$$f(x_{i-1}) = f(x_i) - f'(x_i)h + \frac{f''(x_i)}{2!}h^2 + \frac{f'''(\xi_1)}{3!}h^3,$$

  onde ξ_2 está entre x_{i-1} e x_i.
- Subtraindo as equações acima, temos

$$f(x_{i+1}) - f(x_{i-1}) = 2f'(x_i)h + \frac{f'''(\xi_1)}{3!}h^3 + \frac{f'''(\xi_1)}{3!}h^3.$$

- Resolvendo para $f'(x_i)$, temos

$$f'(x_i) = \frac{f(x_{i+1}) - f(x_{i-1})}{2h} + O(h^2).$$

Diferença finita progressiva com três pontos

- Aproxima $f'(x_i)$ avaliando a função no ponto e nos dois pontos seguintes x_{i+1} e x_{i+2}.
- Aproximação de Taylor em x_{i+1} e x_{i+2},

$$f(x_{i+1}) = f(x_i) + f'(x_i)h + \frac{f''(x_i)}{2!}h^2 + \frac{f'''(\xi_1)}{3!}h^3, \tag{1}$$

$$f(x_{i+2}) = f(x_i) + f'(x_i)2h + \frac{f''(x_i)}{2!}(2h)^2 + \frac{f'''(\xi_2)}{3!}(2h)^3. \tag{2}$$

- Equações 1 e 2 são combinadas de forma que os termos com derivada segunda desapareçam.
- Multiplicando Eq. 1 por 4 e subtraindo Eq. 2, temos

$$4f(x_{i+1}) - f(x_{i+2}) = 3f(x_i) + 2f'(x_i)h + \frac{4f'''(\xi_1)}{3!}h^3 - \frac{f'''(\xi_2)}{3!}(2h)^3.$$

Diferença finita com três pontos

- Resolvendo em $f'(x_i)$, temos

$$f'(x_i) = \frac{-3f(x_i) + 4f(x_{i+1}) - f(x_{i+2})}{2h} + O(h).$$

- Diferença finita regressiva com três pontos

$$f'(x_i) = \frac{f(x_{i-2}) - 4f(x_{i-1}) + 3f(x_i)}{2h} + O(h).$$

Fórmulas de diferenças finitas para a segunda derivada

- Usando as mesmas idéias podemos aproximar a derivada segunda de uma função qualquer por diferenças finitas.
- A derivação das fórmulas são idênticas, porém mais tediosas.
- Fórmula diferença central com três pontos para a derivada segunda

$$f''(x_i) = \frac{f(x_{i-1}) - 2f(x_i) + f(x_{i+1})}{h^2} + O(h^2).$$

- Diferença central com quatro pontos

$$f''(x_i) = \frac{-f(x_{i-2}) + 16f(x_{i-1}) - 30f(x_i) + 16f(x_{i+1}) - f(x_{i+2})}{12h^2} + O(h^4)$$

Fórmulas de diferenças finitas para a segunda derivada

- Diferença progressiva com três pontos

$$f''(x_i) = \frac{f(x_i) - 2f(x_{i+1}) + f(x_{i+2})}{h^2} + O(h).$$

- Diferença regressiva com três pontos

$$f''(x_i) = \frac{f(x_{i-2}) - 2f(x_{i-1}) + f(x_i)}{h^2} + O(h).$$

- Uma infinidade de fórmulas de várias ordens estão disponíveis.
- Fórmulas de diferenciação podem ser obtidas usando polinômios de Lagrange.

Sumário

Erros na diferenciação numérica

- Em todas as fórmulas o erro de truncamento é função de h.
- h é o espaçamento entre os pontos, i.e. $h = x_{i+1} - x_i$.
- Fazendo h pequeno o erro de truncamento será pequeno.
- Em geral usa-se a precisão da máquina, algo como $1e^{-16}$.
- O erro de arredondamento depende da precisão finita de cada computador.
- Mesmo que h possa ser tão pequeno quanto desejado o erro de arredondamento pode crescer quando se diminue h.

Sumário

Extrapolação de Richardson

- Extrapolação de Richardson é usada para obter uma aproximação mais precisa da derivada a partir de duas aproximações menos precisas.
- Considere o valor D de uma derivada (desconhecida) calculada pela fórmula

$$D = D(h) + k_2 h^2 + k_4 h^4, \tag{3}$$

  onde $D(h)$ aproxima D e k_2 e k_4 são termos de erro.
- O uso da mesma fórmula, porém com espaçamento $h/2$ resulta

$$D = D(\frac{h}{2}) + k_2 \left(\frac{h}{2}\right)^2 + k_4 \left(\frac{h}{2}\right)^4. \tag{4}$$

Extrapolação de Richardson

- A Eq. 4 pode ser rescrita (após multiplicar por 4):

$$4D = 4D(\frac{h}{2}) + k_2 h^2 + k_4 \frac{h^4}{4}. \tag{5}$$

- Subtraindo 3 de 5 elimina os termos com h^2 e fornece

$$3D = 4D(\frac{h}{2}) + D(h) - k_4 \frac{3h^4}{4}. \tag{6}$$

- Resolvendo 6, temos

$$D = \frac{1}{3}\left(4D(\frac{h}{2}) + D(h)\right) - k_4 \frac{h^4}{4}. \tag{7}$$

Extrapolação de Richardson

- O erro na Eq. 7 é agora $O(h^4)$. O valor de D é aproximado por

$$D = \frac{1}{3}\left(4D(\frac{h}{2}) + D(h)\right) + O(h^4).$$

- A partir de duas aproximações de ordem inferiores, obtemos uma aproximação de $O(h^4)$ mais precisa.
- Procedimento a partir de duas aproximações com erro $O(h^4)$ mostra que

$$D = \frac{1}{15}\left(16D(\frac{h}{2}) + D(h)\right) + O(h^6).$$

- Aproximação ainda mais precisa.

Exemplo: Extrapolação de Richardson

- Calcule a derivada de $f(x) = \frac{2^x}{x}$ no ponto $x = 2$.
- Solução exata: $\frac{\log(2)2^x}{x} - \frac{2^x}{x^2}$.
- Solução numérica usando diferença central

```
fx <- function(x) (2^x)/x
fpx <- function(x)(log(2)*(2^x))/x - (2^x)/x^2
erro <- fpx(x = 2)/dif_cen(fx = fx, x = 2, h = 0.2)
(erro-1)*100
```

```
# [1] 0.345544
```

- Extrapolação de Richardson

```
D2 <- dif_cen(fx = fx, x = 2, h = 0.2/2)
D <- dif_cen(fx = fx, x = 2, h = 0.2)
der <- (1/3)*( 4*D2 - D)
erro2 <- fpx(x = 2)/der
(erro2-1)*100
```

```
# [1] -0.001585268
```

Sumário

Derivadas parciais

- Para funções com muitas variáveis, a derivada parcial da função em relação a uma das variáveis representa a taxa de variação da função em relação a essa variável, mantendo as demais constantes.
- Assim, as fórmulas de diferenças finitas podem ser usadas no cálculo das derivadas parciais.
- As fórmulas são aplicadas em cada uma das variáveis, mantendo as outras fixas.
- A mesma ideia se aplica para derivadas de mais alta ordem.

Implementação: Derivadas parciais

- Derive $f(\beta_0, \beta_1) = \sum_{i=1}^{n} |y_i - (\beta_0 + \beta_1 x_i)|$.
- Fórmula dois pontos central

```
dif_cen <- function(fx, pt, h, ...) {
  df <- (fx(pt + h, ...) - fx(pt - h, ...))/( (pt + h) - (pt - h))
  return(df)
}
```

- Função a ser diferenciada

```
fx <- function(par, y, x1) {sum ( abs( y - (par[1] + par[2]*x1)) )}
```

- Gradiente usando diferenças finita.

```
grad_fx <- function(fx, par, h, ...) {
  fbeta0 <- function(beta0, beta1, y, x) fx(par = c(beta0, beta1), y = y, x = x)
  fbeta1 <- function(beta1, beta0, y, x) fx(par = c(beta0, beta1), y = y, x = x)
  db0 <- dif_cen(fx = fbeta0, pt = par[1], h = h, beta1 = par[2], y = y, x = x)
  db1 <- dif_cen(fx = fbeta1, pt = par[2], h = h, beta0 = par[1], y = y, x = x)
  return(c(db0, db1))
}
```

Exemplo: Derivadas parciais

- Simulando y_i's e x_i's.

```
set.seed(123)
x <- runif(100)
y <- rnorm(100, mean = 2 + 3*x, sd = 1)
```

- Gradiente numérico

```
grad_fx(fx = fx, par = c(2, 3), h = 0.001, y = y, x1 = x)
```

```
# [1] 6.000000 2.272805
```

- Gradiente analítico

```
c(sum(((y - 2 - 3*x)/abs(y - 2 -3*x))*(-1)),
  sum(((y - 2 - 3*x)/abs(y - 2 -3*x))*(-x)))
```

```
# [1] 6.000000 2.272805
```

Sumário

Uso de funções residentes do R para diferenciação numérica.

- Pacote `numDeriv` implementa derivadas por diferença finita.
- Gradiente

```
require(numDeriv)
args(grad)
```

```
# function (func, x, method = "Richardson", method.args = list(),
#     ...)
# NULL
```

- Hessiano

```
args(hessian)
```

```
# function (func, x, method = "Richardson", method.args = list(),
#     ...)
# NULL
```

Exemplo de aplicação

```r
grad(func = fx, x = c(2, 3), y = y, x1 = x)
```

```
# [1] 6.000000 2.272805
```

```r
hessian(func = fx, x = c(2, 3), y = y, x1 = x)
```

```
#          [,1]     [,2]
# [1,] 58.91271 29.53710
# [2,] 29.53710 48.86648
```