

3. Algoritmo AEIQ-BR

Conforme discutido nos capítulos anteriores, muitos problemas do mundo real podem ser representados através de decisões do tipo combinatórias e numéricas simultaneamente. Por exemplo, seja o processo de modelar um problema de controle, previsão, ou classificação, inevitavelmente se encontrará decisões como: quais variáveis utilizar para modelagem dentre as disponíveis (combinatória); ou como obter os parâmetros dos modelos ou algoritmos utilizados (numérica).

Embora decisões numéricas possam ser representadas através de bits, conforme se observa em Michalewicz [31], esta representação é menos eficiente, pois a dimensionalidade de um problema pode inviabilizar computacionalmente a utilização da maioria dos algoritmos de otimização, conforme Abs da Cruz [2]. Tal limitação levou Abs da Cruz ao desenvolvimento do AEIQ-R com representação numérica do espaço de soluções.

No entanto, pode-se pensar em problemas onde a representação numérica é mais adequada, sem que se desprezem outras decisões de natureza combinatória no cromossoma, cuja representação mais conveniente seria através de genes binários. Para isto, propõe-se o algoritmo AEIQ-BR com representação mista numérica e combinatória de um problema no cromossoma de um algoritmo evolutivo.

Este algoritmo se apresentará como um diferencial para otimização de problemas combinatório e numérico, simultaneamente, herdando ainda as principais características de seus precursores, tais como:

- (1) Capacidade de otimização global de um problema;
- (2) Representação probabilística do espaço de busca, que como consequência irá apresentar alta diversidade populacional em cada indivíduo quântico;
- (3) Necessidade de poucos indivíduos populacionais, e mesmo assim com capacidade de exploração e aproveitamento do espaço de soluções.

3.1. Pseudocódigo

A seguir, os passos do pseudocódigo do algoritmo AEIQ-BR. O detalhamento acontecerá nas seções 3.2 a 3.6.

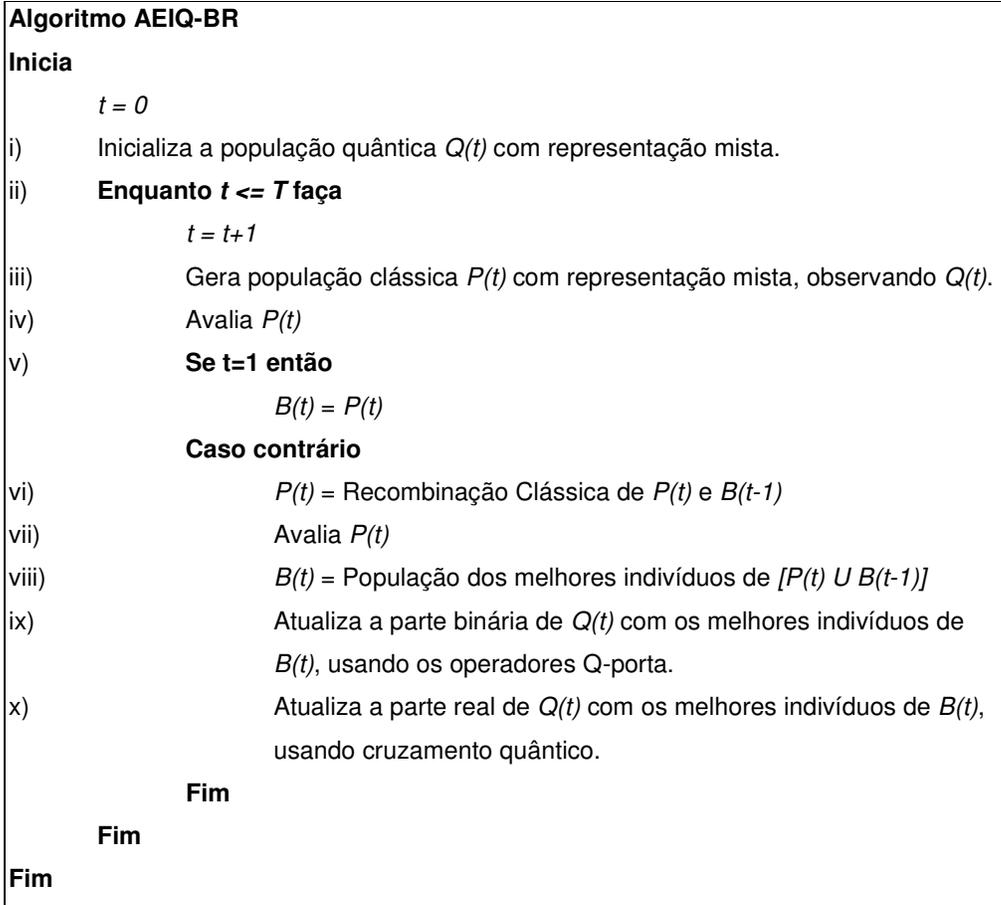


Figura 3.1 – Pseudocódigo do algoritmo AEIQ-BR.

3.2. Representação do Indivíduo Quântico

Na representação de um indivíduo quântico i qualquer num instante de tempo t , através de q_i^t , a representação de seus genes deverá considerar simultaneamente genes quânticos binários e reais, conforme discutido. Sabe-se por Han e Kin [8] que a parte binária de um cromossoma pode ser representada através de um q-bit. Abs da Cruz [2] utilizou o conceito de genes com distribuições probabilísticas para otimização numérica. Embora Abs da Cruz não tenha atribuído um nome específico para estes genes quânticos – cuja utilidade é viabilizar a representação de soluções numéricas no cromossoma – neste trabalho chamar-se-á tais genes de *q-real*. Ou seja, a representação de um gene quântico numérico será feita através de um q-real. Pode-se dizer então que:

$$q_i^t = \left[\left(q_i^t \right)_b \left(q_i^t \right)_r \right]. \quad (3-1)$$

Uma parte do indivíduo será representada através de genes com representação em q-bits, índice b . E outra parte através de genes com representação em q-real, índice r . Sendo assim, a parte binária de um indivíduo quântico com M genes, conforme seção 2.2.1, pode ser escrita da seguinte forma:

$$(q_i^t)_b = \left(g_{i1}^t = \begin{vmatrix} \alpha_{i1}^t \\ \beta_{i1}^t \end{vmatrix}, g_{i2}^t = \begin{vmatrix} \alpha_{i2}^t \\ \beta_{i2}^t \end{vmatrix}, \dots, g_{iM}^t = \begin{vmatrix} \alpha_{iM}^t \\ \beta_{iM}^t \end{vmatrix} \right)_b. \quad (3-2)$$

Da mesma forma, α_{ij}^t e β_{ij}^t uma dupla de números complexos de tal forma que a soma do quadrado de cada um dos elementos, para $j=1, 2, 3, \dots, M$, é igual a 1. E g_{ij}^t trata-se da menor unidade de informação, ou seja, um q-bit. Sendo assim, tais genes serão responsáveis por armazenar preferencialmente soluções para problemas combinatórios, os quais serão representados pelas “observações”³ destes genes quânticos, obtendo-se assim respostas para genes binários do tipo 0 ou 1.

Sobre a parte real do indivíduo quântico, conforme seção 2.2.2 na equação (2-16), tem-se que um gene real pode ser descrito através de uma função distribuição de probabilidade (FDP), integrável em todo seu domínio. Sendo assim, tem-se a seguinte equação:

$$(q_i^t)_r = (g_{i1}^t = p_{i1}^t(x_{i1}^t), g_{i2}^t = p_{i2}^t(x_{i2}^t), g_{i3}^t = p_{i3}^t(x_{i3}^t), \dots, g_{iG}^t = p_{iG}^t(x_{iG}^t))_r. \quad (3-3)$$

Ou seja, a inscrição da parte quântica real de um indivíduo quântico, através de genes representados por funções distribuição de probabilidades. Neste trabalho, a FDP a ser utilizada será a *distribuição uniforme de probabilidade* ou *pulso quadrático*, conforme se observou em Abs da Cruz [2]. Considere uma FDP do tipo uniforme no intervalo a até b para representação de alguma das variáveis x_{ij}^t em (3-3). Tem-se então que:

$$p_{ij}^t(x_{ij}^t) = f(x_{ij}^t; a_{ij}^t; b_{ij}^t) = \begin{cases} \frac{1}{b_{ij}^t - a_{ij}^t}, & \text{se } a_{ij}^t \leq x_{ij}^t \leq b_{ij}^t \\ 0, & \text{caso contrário} \end{cases}. \quad (3-4)$$

A função dada por (3-4) pode ser representada pela figura 3.2. Com base na FDP acima, uma maneira eficiente de representá-la no cromossoma seria através de seus parâmetros a_{ij}^t e b_{ij}^t . No entanto, sua representação acontecerá através do *primeiro momento* e o *domínio* da FPD de x conforme (3-5) e (3-6).

³ A observação de genes quânticos será vista na seção 3.3.

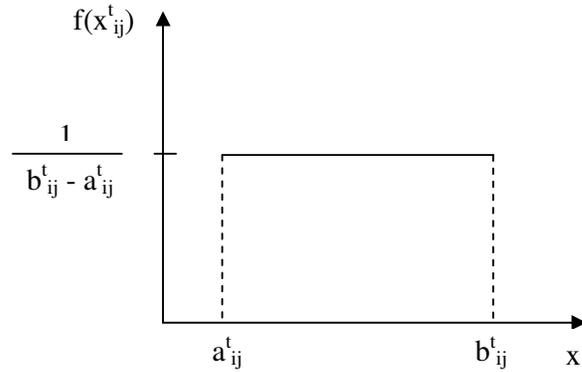


Figura 3.2 – Representação da FDP uniforme no intervalo de a_{ij}^t até b_{ij}^t .

Seja μ_{ij}^t , o primeiro momento da FDP inscrito da seguinte forma:

$$\mu_{ij}^t = E(x_{ij}^t) = \frac{a_{ij}^t + b_{ij}^t}{2}. \quad (3-5)$$

E seja σ_{ij}^t , a amplitude do domínio da FDP da seguinte forma:

$$\sigma_{ij}^t = b_{ij}^t - a_{ij}^t. \quad (3-6)$$

Pode-se reescrever a FDP de (3-4) como:

$$p_{ij}^t(x_{ij}^t) = f(x_{ij}^t; \mu_{ij}^t; \sigma_{ij}^t) = \begin{cases} \frac{1}{\sigma_{ij}^t}, & \text{se } (\mu_{ij}^t - \frac{\sigma_{ij}^t}{2}) \leq x_{ij}^t \leq (\mu_{ij}^t + \frac{\sigma_{ij}^t}{2}), \\ 0, & \text{caso contrário} \end{cases}, \quad (3-7)$$

onde $i=1,2,3,\dots,N$, representando o i -ésimo indivíduo, e $j=1,2,3,\dots,G$ representando o j -ésimo gene q -real. Com isto, pode-se reescrever a parte real do cromossoma de (3-3) através de parâmetros de *centro de variação*, μ_{ij}^t , e *amplitude do domínio*, σ_{ij}^t , como segue:

$$(q_i^t)_r = (g_{i1}^t = p_{i1}^t(x_{ij}^t), g_{i2}^t = p_{i2}^t(x_{ij}^t), g_{i3}^t = p_{i3}^t(x_{ij}^t), \dots, g_{iG}^t = p_{iG}^t(x_{ij}^t))_r = \left(\begin{array}{c|c|c|c|c} \mu_{i1}^t & \mu_{i2}^t & \mu_{i3}^t & \dots & \mu_{iG}^t \\ \sigma_{i1}^t & \sigma_{i2}^t & \sigma_{i3}^t & \dots & \sigma_{iG}^t \end{array} \right)_r. \quad (3-8)$$

A motivação pela representação das FDP conforme parâmetros de centro de variação em (3-5) e amplitude de domínio em (3-6) vem de Abs da Cruz, e a viabilização dos operadores de cruzamento quântico que este utilizou para atualizar os indivíduos quânticos com representação real, conforme detalhado na seção 2.2.2.

Mais adiante, na seção 3.4, se concluirá que esta forma de representação da FDP será útil para se generalizar uma nova heurística de atualização de cada gene

q-real, mais apropriada para nosso problema de otimização, a evolução de uma rede neural. Tem-se então que, um indivíduo quântico, pode ser completamente descrito por:

$$q_i^t = \left[\left(q_i^t \right)_b \left(q_i^t \right)_r \right] = \left[\left(\begin{array}{c|c|c} \alpha_{i1}^t & \alpha_{i2}^t & \dots & \alpha_{iM}^t \\ \beta_{i1}^t & \beta_{i2}^t & & \beta_{iM}^t \end{array} \right)_b \left(\begin{array}{c|c|c|c} \mu_{i1}^t & \mu_{i2}^t & \mu_{i3}^t & \dots & \mu_{iG}^t \\ \sigma_{i1}^t & \sigma_{i2}^t & \sigma_{i3}^t & & \sigma_{iG}^t \end{array} \right)_r \right]. \quad (3-9)$$

E por fim, uma população de N indivíduos quânticos, $i=1,2,3,\dots,N$, conforme *passo i* do pseudocódigo do algoritmo, figura 3.1, pode ser inscrita como:

$$Q(t) = [q_1^t, q_2^t, q_3^t, \dots, q_N^t]. \quad (3-10)$$

A inicialização de $Q(t)$ no passo i do algoritmo acontecerá de tal forma que *a probabilidade de se obter uma solução qualquer em todo o espaço de busca é única e igual para todos os indivíduos*. Isto somente acontecerá quando para cada q-bit, a probabilidade de se observar um estado for $|0\rangle = |1\rangle = 1/2$, e a probabilidade de se observar um gene q-real for dado pelos parâmetros de centro e variação originais da FDP para algum x .

Isto equivale a dizer que na inicialização do algoritmo ocorre uma *superposição* de todos os possíveis estados, com igual probabilidade de observação de todos os estados. Isto pode ser entendido com uma característica de *diversidade*, representada em um único indivíduo quântico, contrariamente aos algoritmos evolutivos tradicionais onde se requer a presença de diversos indivíduos clássicos.

Por exemplo, seja um exemplo fictício de representação de uma população quântica para um problema. Suponha uma população $Q(t)$ no primeiro passo do algoritmo, com somente dois indivíduos q_1^t e q_2^t . Suponha que as partes quânticas binária destes cromossomas tenham um total de 4 genes, cada gene representando uma decisão de natureza dicotômica na decodificação, ou combinatória. Suponha ainda que a *parte quântica real* deste cromossoma tenha 4 genes, cada gene responsável por representar uma decisão de natureza numérica na decodificação, para variáveis que estão todas definidas no domínio $\{-4,4\}$. Sendo assim, tal população pode ser descrita conforme (3-11).

E assim, em cada gene quântico da parte binária de cada indivíduo, ou seja, em cada q-bit, a probabilidade de se obter os estados 0 ou 1 são iguais e dadas por $1/2$. Onde $1/2$ é obtido elevando-se qualquer α_{ij}^t ou β_{ij}^t ao quadrado, com $i=1,2$ e $j=1,2,3,4$. Observe também que a probabilidade de cada q-real, dada pela FDP (3-7) de parâmetros μ_{ij}^t e σ_{ij}^t , é igual para qualquer valor numérico no domínio da

$f(x_{ij}^t)$, para qualquer $i=1,2$ e $j=1,2,3,4$, pois a FDP utilizada é a uniforme. Como dito anteriormente, isto significa dizer que um único indivíduo é capaz de representar no passo inicial uma superposição de todas as possíveis soluções do domínio do problema.

$$\begin{aligned}
 Q(t) = & \\
 \{q_1^t, q_2^t\} = & \\
 \left\{ \left[(q_1^t)_b, (q_1^t)_r \right], \left[(q_2^t)_b, (q_2^t)_r \right] \right\} = & \\
 \left\{ \left[\left(\alpha_{11}^t; \alpha_{12}^t; \alpha_{13}^t; \alpha_{14}^t \right) \left(\mu_{11}^t; \mu_{12}^t; \mu_{13}^t; \mu_{14}^t \right) \right] \left[\left(\alpha_{21}^t; \alpha_{22}^t; \alpha_{23}^t; \alpha_{24}^t \right) \left(\mu_{21}^t; \mu_{22}^t; \mu_{23}^t; \mu_{24}^t \right) \right] \right\} = & \\
 \left\{ \left[\left(\beta_{11}^t; \beta_{12}^t; \beta_{13}^t; \beta_{14}^t \right)_b \left(\sigma_{11}^t; \sigma_{12}^t; \sigma_{13}^t; \sigma_{14}^t \right)_r \right] \left[\left(\beta_{21}^t; \beta_{22}^t; \beta_{23}^t; \beta_{24}^t \right)_b \left(\sigma_{21}^t; \sigma_{22}^t; \sigma_{23}^t; \sigma_{24}^t \right)_r \right] \right\} = & \\
 \left\{ \left[\left(\frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}} \right) \left(0; 0; 0; 0 \right) \right] \left[\left(\frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}} \right) \left(0; 0; 0; 0 \right) \right] \right\} = & \\
 \left\{ \left[\left(\frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}} \right)_b \left(8; 8; 8; 8 \right)_r \right] \left[\left(\frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}} \right)_b \left(8; 8; 8; 8 \right)_r \right] \right\} &
 \end{aligned}
 \tag{3-11}$$

Tendo em vista os resultados acima, pode-se compreender o primeiro passo do algoritmo AEIQ-BR. A partir do segundo passo o algoritmo iniciará a evolução e assim se manterá até que atinja um número máximo de instantes de tempo, ou como é chamado mais frequentemente no contexto de computação evolutiva, número de máximo de *gerações*.

Antes de se continuar com o passo a passo de como é gerada uma população clássica $P(t)$ com a *observação de* $Q(t)$, introduzir-se-á também as formas de representação a serem utilizadas para genes e indivíduos clássicos. Sendo assim, defini-se um indivíduo clássico da seguinte forma:

$$c_i^t = \left[\left(c_i^t \right)_b, \left(c_i^t \right)_r \right] = \left[\left(x_{i1}^t, x_{i2}^t, x_{i3}^t, \dots, x_{iM}^t \right)_b, \left(x_{i1}^t, x_{i2}^t, x_{i3}^t, \dots, x_{iG}^t \right)_r \right]. \tag{3-12}$$

O índice i para o i -ésimo indivíduo clássico, com $i=1,2,3,\dots,N$, o *tamanho da população*. O índice $j=1,2,3,\dots,M$, o *total de genes clássicos binários*. O índice $h=1,2,3,\dots,G$, o *total de genes clássicos reais*. Com variáveis $(x_{ij}^t)_b$ assumindo somente valores 0 ou 1. E variáveis $(x_{ij}^t)_r$ assumindo valores no domínio da FDP em (3-7). E por fim, um instante de tempo, ou geração, t qualquer.

Sendo assim, uma população de N indivíduos c_i^t pode ser escrita da seguinte forma:

$$C(t) = \left[c_1^t, c_2^t, c_3^t, \dots, c_N^t \right]. \tag{3-13}$$

A seguir, o detalhamento do processo de geração da população clássica $P(t)$ a partir de $Q(t)$.

3.3. Observação do Indivíduo Quântico

É através da observação de indivíduos quânticos que soluções manipuláveis são obtidas para o problema em modelagem, e assim, não se manipulará mais somente representações probabilísticas do espaço de soluções. Em outras palavras, somente com a observação dos indivíduos quânticos que se obtêm parâmetros suficientes para avaliar, ou simular, como os resultados se comportam em termos de um objetivo de otimização.

A observação do indivíduo quântico é feita através de um processo de seleção aleatória com iguais probabilidades de uma solução no espaço de busca representado por cada cromossoma. Suponha um gene q-bit qualquer $(g^t_{ij})_b$ de um indivíduo q_i^t qualquer. Graficamente, para a observação deste q-bit $(g^t_{ij})_b$ em um passo inicial t inicial, onde $(\alpha^t_{ij})^2 + (\beta^t_{ij})^2 = 1$, observa-se o seguinte:

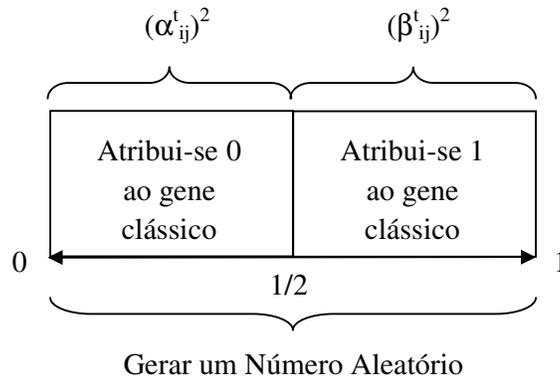


Figura 3.3 – Representação do processo de observação de um q-bit $(g^t_{ij})_b$.

Tratando-se da obtenção dos M genes binários de um total de N_c indivíduos clássicos observando-se os q-bits de N_Q indivíduos quânticos, o procedimento descrito na figura 3.4 é executado.

Tanto a figura 3.3 quanto a figura 3.4, representam procedimentos que ocorrerão em um passo inicial do algoritmo AEIQ-BR. Diz-se passo inicial, pois em passos seguintes ocorrerão atualizações⁴ nos indivíduos quânticos, e as probabilidades de se observar estados 0 e 1 serão desiguais, porém sempre respeitando a

⁴ Detalhamento a seguir na seção 3.4.

soma total igual a 1. Sendo assim, este processo de geração de números aleatórios e atribuição de genes clássicos 0 ou 1 seria executado para cada um dos q-bits e suas respectivas probabilidades, de todos os indivíduos quânticos.

Observação dos Q-bits

Inicia

Para i=1,2,3,...,N_Q-número de indivíduos quânticos

Seleciona-se q_i^t , o i-ésimo indivíduo quântico em $Q(t)$.

Para j=1,2,3,...,M-número de q-bits presentes em q_i^t

Calcula-se a probabilidade de se ter genes 0 ou 1 dado, respectivamente, por $(\alpha_{ij}^t)^2$ e $(\beta_{ij}^t)^2$, em $(g_{ij}^t)_b$.

Gera-se um número aleatório entre 0 e 1.

Se aleatório entre 0 e uma probabilidade de crossover clássico C_c .

Atribui-se ao gene clássico $(x_{ij}^t)_b$ o valor 0.

Caso contrário

Atribui-se ao gene clássico $(x_{ij}^t)_b$ o valor 1.
(crossover tipo uniforme)

Fim Se

Fim Para

Fim Para

Fim

Figura 3.4 – Pseudocódigo da observação da parte binária de um indivíduo quântico.

Já a geração de um gene clássico real $(x_{ij}^t)_r$ qualquer, através da observação de um q-real $(g_{ij}^t)_r$ qualquer, é feita através da *inversa da distribuição acumulada de probabilidade* de cada um dos genes quânticos, conforme (3-14) a (3-16). Como observado em (3-7), um gene quântico real, ou q-real, pode ser representado por uma FDP do tipo $p_{ij}^t(x_{ij}^t) = f_{ij}^t(x_{ij}^t; \mu_{ij}^t; \sigma_{ij}^t)$. Integrando esta função no domínio de $-\infty$ a x_{ij}^t , pode-se obter a seguinte equação:

$$P_{ij}^t(x_{ij}^t) = F(x_{ij}^t; \mu_{ij}^t; \sigma_{ij}^t) = \begin{cases} 0 & , \text{ se } x_{ij}^t < (\mu_{ij}^t - \frac{\sigma_{ij}^t}{2}) \\ \frac{x_{ij}^t - \mu_{ij}^t}{\sigma_{ij}^t} + \frac{1}{2}, & \text{ se } (\mu_{ij}^t - \frac{\sigma_{ij}^t}{2}) \leq x_{ij}^t \leq (\mu_{ij}^t + \frac{\sigma_{ij}^t}{2}) \\ 1 & , \text{ se } x_{ij}^t \geq (\mu_{ij}^t + \frac{\sigma_{ij}^t}{2}) \end{cases} \quad (3-14)$$

É importante mencionar que uma função cumulativa de probabilidade necessariamente apresentará uma imagem no intervalo de 0 a 1. Para facilitar a

compreensão, representar-se-á pela letra y , o resultado da equação em (3-14), ou seja:

$$y_{ij}^t = P_{ij}^t(x_{ij}^t) = F(x_{ij}^t; \mu_{ij}^t; \sigma_{ij}^t). \quad (3-15)$$

Desta forma, a inversa da função distribuição cumulativa de probabilidade em (3-14), definida no domínio 0 até 1, poderá ser escrita da seguinte forma:

$$\left(x_{ij}^t\right)_r = P_{ij}^{t-1}(y_{ij}^t) = F^{-1}(y_{ij}^t; \mu_{ij}^t; \sigma_{ij}^t) = \begin{cases} 0 & , \text{ se } y_{ij}^t < 0 \\ \sigma_{ij}^t \bullet \left(y_{ij}^t - \frac{1}{2}\right) + \mu_{ij}^t, & \text{ se } 0 \leq y_{ij}^t \leq 1 \\ 1 & , \text{ se } y_{ij}^t \geq 1 \end{cases} \quad (3-16)$$

Tendo em mãos a função acima, é possível observar os q-reals, e obter-se valores numéricos para os indivíduos clássicos. Isto também será feito através de um processo de geração de números aleatórios no intervalo de 0 a 1, ou seja, no domínio da função em (3-16), e posterior observação de $(x_{ij}^t)_r$. Tem-se então:

Observação dos Q-reals

Inicia

Para $i=1,2,3,\dots, N_Q$ -número de indivíduos quânticos
 Seleciona-se q_i^t , o i -ésimo indivíduo quântico em $Q(t)$.

Para $j=1,2,3,\dots, G$ -número de q-reals presentes em q_i^t
 Gera-se um número aleatório entre 0 e 1.
 Atribuir a y_{ij}^t o valor deste número aleatório.
 Obtenha $(x_{ij}^t)_r$, através da inversa da função distribuição cumulativa de probabilidade em (3-16).

Fim Para

Fim Para

Fim

Figura 3.5 – Pseudocódigo da observação da parte real de um indivíduo quântico.

Suponha a mesma população quântica conforme (3.11). Suponha ainda um vetor de números aleatórios no intervalo de 0 até 1. Considerando os q-bits definidos por (3-11), as probabilidades de se observar um gene clássico binário 0 ou 1 são iguais a $\frac{1}{2}$. Sendo assim, através do procedimento descrito na figura 3.5, valores observados são gerados para genes clássicos conforme tabela 3.1.

Por outro lado, para valores observados a partir dos q-reals, o que se faz é introduzir os números aleatórios obtidos na inversa acumulada da FDP em (3-16), com parâmetros de centro iguais a 0 e amplitude de variação igual a 8, para cada x_{ij}^t , com $i=1,2$ e $j=1,2,3,4$. Sendo assim, através de um simples cálculo matemático

Gene	Número Aleatório	Genes Clássicos
$(q_{11})^t_b$	0,723	1
$(q_{12})^t_b$	0,006	0
$(q_{13})^t_b$	0,586	1
$(q_{14})^t_b$	0,166	0
$(q_{11})^t_r$	0,828	2,627
$(q_{12})^t_r$	0,174	-2,604
$(q_{13})^t_r$	0,306	-1,551
$(q_{14})^t_r$	0,518	0,146
$(q_{21})^t_b$	0,297	0
$(q_{22})^t_b$	0,056	0
$(q_{23})^t_b$	0,365	0
$(q_{24})^t_b$	0,567	1
$(q_{21})^t_r$	0,647	1,174
$(q_{22})^t_r$	0,491	-0,071
$(q_{23})^t_r$	0,913	3,304
$(q_{24})^t_r$	0,610	0,882

Tabela 3.1 – Exemplo de obtenção das probabilidades de um indivíduo quântico.

introduzindo-se cada número aleatório gerado na equação (3-17) a seguir, pode-se obter o gene clássico numérico, os quais também estarão representados na tabela 3.1. É importante notar na equação anterior que quando y^t_{ij} for igual a 1, $(x^t_{ij})_r$ será máximo igual a 4, enquanto que se y^t_{ij} for igual a 0, $(x^t_{ij})_r$ será mínimo igual a -4.

$$(x^t_{ij})_r = P^{t-1}_{ij}(y^t_{ij}) = F^{-1}(y^t_{ij}; 0; 8) = \begin{cases} 0 & , \text{ se } y^t_{ij} < 0 \\ 8 \cdot (y^t_{ij} - \frac{1}{2}), & \text{ se } 0 \leq y^t_{ij} \leq 1. \\ 1 & , \text{ se } y^t_{ij} \geq 1 \end{cases} \quad (3-17)$$

Tendo em vista os resultados anteriores desta seção, obtêm-se todo o conhecimento necessário para compreensão do processo de *observação de um gene quântico*, seja a parte binária q-bit ou numérica q-real, e, conseqüentemente, obtenção dos genes clássicos conforme *passo iii* da figura 3.1.

Faz-se uma ressalva somente ao número de indivíduos clássicos que podem ser observados, e o número de indivíduos quânticos que se dispõe, os quais não necessariamente devem ser iguais. O programador encarregado de otimizar um problema através do algoritmo aqui proposto, poderá utilizar-se de uma *estratégia evolutiva de observação de mais de um indivíduo clássico* pertencentes a P(t), *através do mesmo indivíduo quântico* pertencente a Q(t), conforme passo iii do algoritmo AEIQ-BR na figura 3.1.

Para isso, espera-se que o número de indivíduos clássicos seja sempre um múltiplo do número de indivíduos quânticos, pois o algoritmo não dispõe de nenhuma heurística para privilegiar qualquer quântico na geração de mais clássicos que os demais. Sendo assim, é necessário que o procedimento detalhado nas figuras 3.4 e 3.5 sejam executados um número z de vezes, onde estariam sendo observado um total de $z \times N_Q$ indivíduos clássicos.

Conforme discutido no capítulo 2, a observação de um estado quântico pode levar ao colapso de todos os outros possíveis estados. Porém isto acontece em um computador quântico. Deve-se lembrar que este algoritmo busca inspiração quântica e que, portanto, será executado em um computador clássico. Daí a solução técnica para observação mais de uma vez cada indivíduo quântico.

Esta característica estará intrinsecamente associada à capacidade do algoritmo de *aproveitamento*, mencionada no capítulo introdutório desta dissertação, e frequentemente mencionada por Han e Kin [4] e Abs da Cruz [2] em seus trabalhos, conforme discutido nas seções 2.1.1 e 2.1.2. Isto porque na prática se estaria dando uma maior oportunidade para o algoritmo de vasculhar melhor certa região do espaço de soluções representada por um único indivíduo quântico, uma vez que diversos indivíduos clássicos estariam sendo gerados do mesmo. Já a capacidade de *exploração* – também mencionada nos trabalhos de Han e Kin e Abs da Cruz – dar-se em função da capacidade de superposição de estados de um indivíduo quântico, uma vez que indivíduos clássicos podem ser gerados em diversos extremos de um domínio do espaço de soluções.

No algoritmo AEIQ-BR, parâmetros como número total de indivíduos quânticos na população e número total de indivíduos clássicos a serem observados, serão denotados respectivamente pelas siglas N_Q e N_C . Em caso de N_C ser menor que N_Q , alguns indivíduos quânticos não estariam sendo observados, e, portanto, características como exploração e aproveitamento poderiam estar comprometidas.

Um detalhe importante a ser acrescentado é que o algoritmo não dispõe de inteligência para julgar durante a evolução, quais indivíduos quânticos são mais promissores para otimização do problema.

Sendo assim é recomendado que N_C seja um múltiplo de N_Q , pois assim cada indivíduo quântico teria condições iguais de se gerar, através da observação, a mesma quantidade de indivíduos clássicos.

Por outro lado, existem heurísticas para se determinar quais indivíduos clássicos são mais ou menos promissores, característica esta presente em todo algoritmo evolutivo, sendo determinado de acordo com as particularidades do problema que se deseja otimizar. Tais heurísticas serão úteis para os passos iv e vii do pseudocódigo na figura 3.1, sendo discutidas mais profundamente em 3.7.3, quando da definição de funções de aptidão para uma aplicação real de otimização do algoritmo AEIQ-BR.

A atualização de um indivíduo quântico – conforme será discutido na seção a seguir – tem como objetivo fazer com que cada indivíduo represente melhor regiões mais promissoras de soluções para o problema, sem, portanto, fazer alguma comparação de melhor ou pior representação entre indivíduos quânticos.

É importante dizer que o *número máximo de passos ou gerações* do algoritmo AEIQ-BR, onde ocorrerão processos de representação e observação de um indivíduo quântico discutidos até então, será determinado manualmente na configuração inicial do algoritmo (critério de parada), e no pseudocódigo é representado pela letra t conforme figura 3.1.

3.4. Atualização do Indivíduo Quântico

Observou-se na seção anterior que parâmetros iniciais para q-bits e q-reals são configurados de tal forma que estes representem a superposição de todos os possíveis estados. Em outras palavras, a probabilidade de observação de qualquer indivíduo clássico, através de qualquer indivíduo quântico é igual para todas as soluções no domínio do problema, no passo inicial do algoritmo.

Para que o algoritmo tenha alguma vantagem evolutiva, é recomendável que os parâmetros q-bits e q-reals sejam ajustados de tal forma que estes passarão a representar em um passo posterior – outra geração – espaços de possíveis soluções mais promissoras para o problema, associadas a ganhos mensurados através de uma função de avaliação em iv e vii da figura 3.1, e os indivíduos clássicos obtidos até então.

Destaca-se o seguinte trecho do pseudocódigo: “*atualiza a parte binária de $Q(t)$ com os melhores indivíduos de $B(t)$* ”. Funcionalmente, os melhores indivíduos de $B(t)$ fornecerão as direções de atualização, para que os indivíduos em $Q(t)$ representem melhores soluções, as quais em gerações seguintes podem ser mais

promissoras. Tecnicamente, na atualização de um q_i^t através de um c_j^t , para quaisquer i e j , uma análise comparativa entre q-bits e seus correspondentes genes clássicos binários, e q-reals e seus correspondentes genes clássicos reais será feita. Esta atualização tem o propósito de fazer com que q_i^t represente melhor c_j^t , empiricamente uma solução promissora, já que c_j^t pertencerá a $B(t)$, conforme passos *viii*, *ix*, *x* da figura 3.1.

Seja um gene $(g_{ij}^t)_b$, um q-bit qualquer pertencente a um indivíduo q_i^t . Seja também $(x_{ij}^t)_b$, um gene clássico qualquer pertencente a um indivíduo c_i^t . Se através de alguma heurística se comprovasse que indivíduos observados c_i^t têm um melhor desempenho quando o gene clássico observado $(x_{ij}^t)_b$, do correspondente quântico $(g_{ij}^t)_b$, é igual a 1, a probabilidade $(\beta_{ij}^t)^2$ deste q-bit deveria ser *aumentada*. Desta forma, seria mais provável que no futuro este gene $(g_{ij}^t)_b$ gerasse com uma maior chance um gene clássico correspondente $(x_{ij}^t)_b$ igual a 1.

E contrariamente, se através de alguma heurística se comprovasse que indivíduos observados c_i^t têm um melhor desempenho quando o gene clássico observado $(x_{ij}^t)_b$, do correspondente quântico $(g_{ij}^t)_b$, é igual a 0, a probabilidade $(\beta_{ij}^t)^2$ deste q-bit deveria ser *diminuída*.

E assim, através de um processo contínuo de atualizações do gene $(g_{ij}^t)_b$ a cada geração, este provavelmente convergiria para uma probabilidade máxima igual a 1 de se observar um gene clássico igual a 1, e um probabilidade mínima, e, portanto, impossível de se observar um gene clássico igual a 0.

A responsabilidade de se atualizar qualquer q-bit é atribuída à porta quântica de rotação $U(\theta)$, como visto na seção 2.1.1, e no pseudocódigo da figura 3.1 será executada no passo *ix*. Seja um gene $(g_{ij}^t)_b$ de um q-bit qualquer, com $i=1,2,3,\dots,N$, total de indivíduos, e $j=1,2,3,\dots,M$, total de q-bits possíveis, conforme (3-2). Seja também $(x_{ij}^t)_b$, o gene clássico correspondente em um indivíduo clássico qualquer, pertencente à população dos melhores indivíduos $B(t)$. Assim a atualização $(g_{ij}^t)_b$ pode ser obtida por:

$$\left(g_{ij}^{t+1} \right)_b = \begin{bmatrix} \alpha_{ij}^{t+1} \\ \beta_{ij}^{t+1} \end{bmatrix}_b = \begin{cases} \begin{bmatrix} \cos(\theta) & \text{seno}(\theta) \\ -\text{seno}(\theta) & \cos(\theta) \end{bmatrix} \bullet \begin{bmatrix} \alpha_{ij}^t \\ \beta_{ij}^t \end{bmatrix}, & \text{se } (x_{ij}^t)_b = 0 \\ \begin{bmatrix} \cos(\theta) & -\text{seno}(\theta) \\ \text{seno}(\theta) & \cos(\theta) \end{bmatrix} \bullet \begin{bmatrix} \alpha_{ij}^t \\ \beta_{ij}^t \end{bmatrix}, & \text{se } (x_{ij}^t)_b = 1 \end{cases}, \quad (3-18)$$

onde θ um ângulo de rotação definido pelo usuário. Esta porta quântica é a mesma usada por Han e Kin [8] e foi amplamente discutida no capítulo 2.2.1.

Raciocínio análogo pode ser desenvolvido para um gene $(g^t_{ij})_r$ agora representando um q-real qualquer. Seja uma população de melhores indivíduos clássicos $B(t)$, numa análise comparativa do gene correspondente $(x^t_{ij})_r$ de algum destes indivíduos clássicos com o gene $(g^t_{ij})_r$, é provável que este represente um centro de variação mais promissor para o problema – pois este é um dos melhores indivíduos de $B(t)$ –, e, portanto, μ^t_{ij} deve ser atualizado na direção de $(x^t_{ij})_r$. Da mesma forma para σ^t_{ij} , pois se $(x^t_{ij})_r$ representa uma solução numa região mais promissora, as buscas ao redor desta região deve ser intensificada, ou seja, *aproveitada*.

A atualização de μ^t_{ij} pode ser observada em Abs da Cruz [2] e é feita conforme (3-19). Considere agora $(g^t_{ij})_r$ um q-real qualquer, com $i=1,2,3,\dots,N$, total de indivíduos, e $j=1,2,3,\dots,G$, total de q-reals possíveis, conforme (3-3). Seja também $(x^t_{ij})_r$, o gene clássico correspondente em um indivíduo clássico qualquer, pertencente à população dos melhores indivíduos $B(t)$. A atualização μ_{ij} pode ser obtida da seguinte forma, conforme Abs da Cruz [2]:

$$\mu^{t+1}_{ij} = \mu^t_{ij} + \lambda^t ((x^t_{ij})_r - \mu^t_{ij}), \quad (3-19)$$

onde λ^t é um número aleatório entre 0 e 1, gerado a cada geração t .

Observe que se $\lambda^t=1$, μ^{t+1}_{ij} será exatamente igual a $(x^t_{ij})_r$. Enquanto que se $\lambda^t=0$, μ^{t+1}_{ij} será exatamente igual a μ^t_{ij} . Ou seja, a atualização do centro μ^t_{ij} de um $(g^t_{ij})_r$, na direção de um gene clássico $(x^t_{ij})_r$ dada a uma taxa λ^t .

Já a atualização do σ^t_{ij} – diferentemente da que pode ser encontrado em Abs da Cruz [2] – é feita através de um *decaimento exponencial* com velocidade de atualização dada pela amplitude presente em $(x^t_{ij})_r$, *de todos os indivíduos na população $B(t)$* . Isto pode ser resumido pela seguinte equação:

$$\sigma^{t+1}_{ij} = \sigma^t_{ij} + \lambda^t \cdot ((\max((x^t_{ij})_r) - \min((x^t_{ij})_r)) - \sigma^t_{ij}); \quad (3-20)$$

com σ^t_{ij} fixo para algum indivíduo quântico i e q-real j ; $\max((x^t_{ij})_r)$ o máximo o gene clássico real *entre todos os melhores indivíduos de $B(t)$* , ou seja, entre todos os x^t_{ij} mantendo σ^t_{ij} fixo; $\min((x^t_{ij})_r)$ o mínimo do gene clássico real ij , também *entre todos os melhores indivíduos de $B(t)$* , entre todos os x^t_{ij} mantendo σ^t_{ij} fixo; e λ^t o mesmo número aleatório entre 0 e 1 utilizado em (3-19), determinando a velocidade de atualização.

Para fins de exemplificação, suponha λ^t constante entre todas as gerações t , igual a um número fixo 0,9. Suponha também que a cada geração t , $\max((x^t_{ij})_r)$ e $\min((x^t_{ij})_r)$ sejam dados a uma proporção aproximada, respectivamente, de 0,99 e 0,01 da amplitude de σ^t_{ij} . Suponha também a amplitude de um gene $\sigma^{t=1}_{ij}$ no primeiro passo do algoritmo com valor igual a 4, ou seja, $\sigma^{t=1}_{ij}=4$. Pode se observar que o decaimento exponencial conforme mencionado pode ser dado pela figura 3.6.

Observe que $\sigma^{t=1}_{ij}=4$ decai assintoticamente sobre eixo x , porém sem nunca assumir valor nulo. Esta forma de se atualizar o parâmetro σ^t_{ij} foi generalizada para este trabalho, com base na heurística de sua aplicação conforme será discutido na seção 3.7.

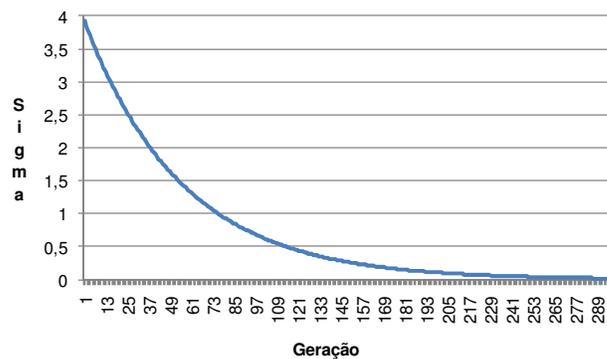


Figura 3.6 – Exemplo de decaimento do parâmetro σ de um indivíduo quântico.

Pesquisando autores em busca de menções a este tipo de atualização encontrou-se Abs da Cruz [2], o qual considera como válida alguma heurística de atualização dos parâmetros σ^t_{ij} da forma *decaimento exponencial*, tendo ele se utilizado de um processo similar, a regra do 1/5 conforme Michalewicz (pág. 94).

O passo a passo para atualização de todos os indivíduos quânticos i , $i=1,2,3,\dots,N_Q$, em todos os q-bits j , $j=1,2,3,\dots,M$, e q-reals h , $h=1,2,3,\dots,G$, no passo ix e x do algoritmo AEIQ-BR, em cada geração t é dado pela figura 3.7.

Como N_C – número total de clássicos – poderá ser maior que N_Q – número total de quânticos –, selecionam-se por ordem decrescente de aptidão (o melhor indivíduo é selecionado primeiro) indivíduos clássicos de $B(t)$, sem repetição. A probabilidade de cruzamento quântico C_Q permanecerá constante durante todas as gerações.

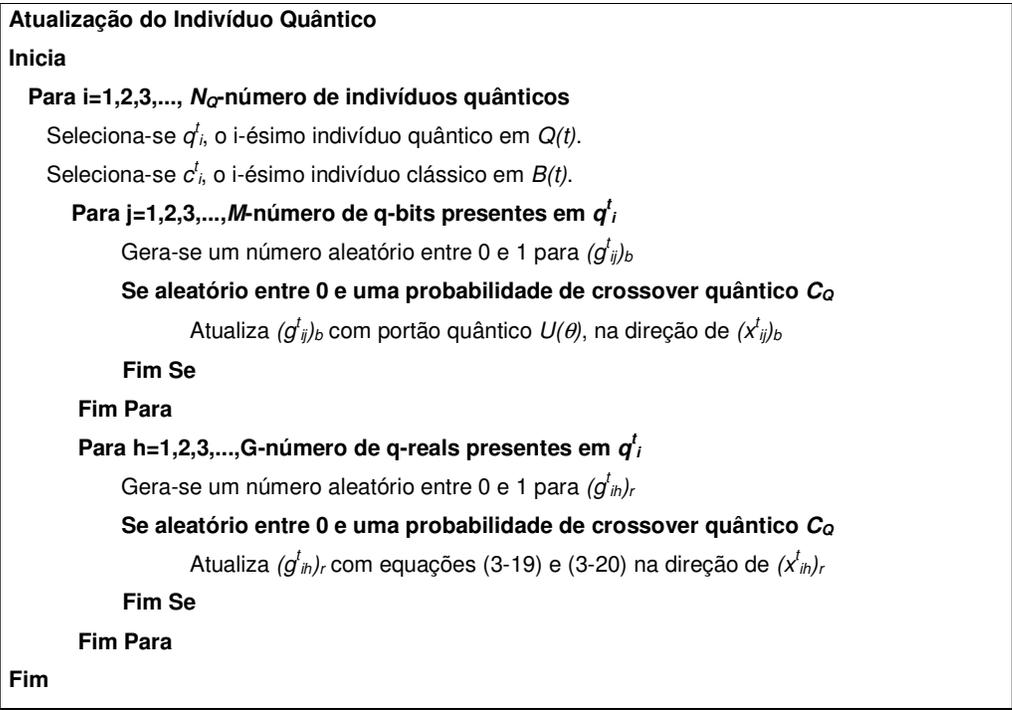


Figura 3.7 – Pseudocódigo da atualização de um indivíduo quântico.

Índice	Genes Quânticos				Genes Clássicos
	α	β	μ	σ	
$(q_{11})_b^t / (x_{11})_b^t$	$1/\sqrt{2}$	$1/\sqrt{2}$	-	-	1
$(q_{12})_b^t / (x_{12})_b^t$	$1/\sqrt{2}$	$1/\sqrt{2}$	-	-	0
$(q_{13})_b^t / (x_{13})_b^t$	$1/\sqrt{2}$	$1/\sqrt{2}$	-	-	1
$(q_{14})_b^t / (x_{14})_b^t$	$1/\sqrt{2}$	$1/\sqrt{2}$	-	-	0
$(q_{11})_r^t / (x_{11})_r^t$	-	-	0	8	2,928
$(q_{12})_r^t / (x_{12})_r^t$	-	-	0	8	3,941
$(q_{13})_r^t / (x_{13})_r^t$	-	-	0	8	-0,298
$(q_{14})_r^t / (x_{14})_r^t$	-	-	0	8	-0,379
$(q_{21})_b^t / (x_{21})_b^t$	$1/\sqrt{2}$	$1/\sqrt{2}$	-	-	0
$(q_{22})_b^t / (x_{22})_b^t$	$1/\sqrt{2}$	$1/\sqrt{2}$	-	-	0
$(q_{23})_b^t / (x_{23})_b^t$	$1/\sqrt{2}$	$1/\sqrt{2}$	-	-	1
$(q_{24})_b^t / (x_{24})_b^t$	$1/\sqrt{2}$	$1/\sqrt{2}$	-	-	1
$(q_{21})_r^t / (x_{21})_r^t$	-	-	0	8	3,443
$(q_{22})_r^t / (x_{22})_r^t$	-	-	0	8	1,688
$(q_{23})_r^t / (x_{23})_r^t$	-	-	0	8	0,478
$(q_{24})_r^t / (x_{24})_r^t$	-	-	0	8	-0,425

Tabela 3.2 – Exemplo de atualização de uma população quântica $Q(t)$.

Resumidamente, o que acontecerá através de um processo evolutivo será a convergência de probabilidade máxima igual a 1 de se observar um $(g^t_{ij})_b$ igual a 1 ou 0, para $i=1,2,3,\dots,N$, e $j=1,2,3,\dots,M$. E amplitude de variação σ^t_{ij} praticamente nula ao redor de um centro μ^t_{ij} , para $i=1,2,3,\dots,N$, e $j=1,2,3,\dots,G$. Ou seja, a diversidade presente nos indivíduos quânticos se perderá a cada geração, e ao final, cada $(g^t_{ij})_b$ e $(g^t_{ij})_r$ convergirão para um único valor, a melhor solução encontrada para o problema a ser otimizado.

Suponha uma população quântica conforme dado por (3-11). Os parâmetros que definem q-bits e q-reals podem ser observados na tabela 3.2. Suponha também dois indivíduos clássicos quaisquer de $B(t)$ selecionados para atualizar a população quântica conforme mesma tabela 3.2. Sendo assim, a atualização dos indivíduos quânticos utilizando as equações dadas em (3-18) com $\theta=0,05*PI$ (conforme recomendações em Han e Kin [4]), (3-19) e (3-20) com $\lambda^t=0,281$ (obtido aleatoriamente), para um instante de tempo t qualquer, geraria os seguintes valores:

Índice	Genes Quânticos			
	α	β	μ	σ
$(q_{11})^t_b / (x_{11})^t_b$	0,572	0,416	-	-
$(q_{12})^t_b / (x_{12})^t_b$	0,416	0,572	-	-
$(q_{13})^t_b / (x_{13})^t_b$	0,572	0,416	-	-
$(q_{14})^t_b / (x_{14})^t_b$	0,416	0,572	-	-
$(q_{11})^t_r / (x_{11})^t_b$	-	-	0,823	5,897
$(q_{12})^t_r / (x_{12})^t_b$	-	-	1,107	6,385
$(q_{13})^t_r / (x_{13})^t_b$	-	-	-0,084	5,970
$(q_{14})^t_r / (x_{14})^t_b$	-	-	-0,106	5,765
$(q_{21})^t_b / (x_{21})^t_b$	0,416	0,572	-	-
$(q_{22})^t_b / (x_{22})^t_b$	0,572	0,416	-	-
$(q_{23})^t_b / (x_{23})^t_b$	0,416	0,572	-	-
$(q_{24})^t_b / (x_{24})^t_b$	0,572	0,416	-	-
$(q_{21})^t_r / (x_{21})^t_b$	-	-	0,967	5,897
$(q_{22})^t_r / (x_{22})^t_b$	-	-	0,474	6,385
$(q_{23})^t_r / (x_{23})^t_b$	-	-	0,134	5,970
$(q_{24})^t_r / (x_{24})^t_b$	-	-	-0,119	5,765

Tabela 3.3 – População quântica atualizada.

A figura 3.8 resume de forma gráfica a atualização que um q^t_{ij} qualquer recebe na forma de representação de um espaço de soluções.

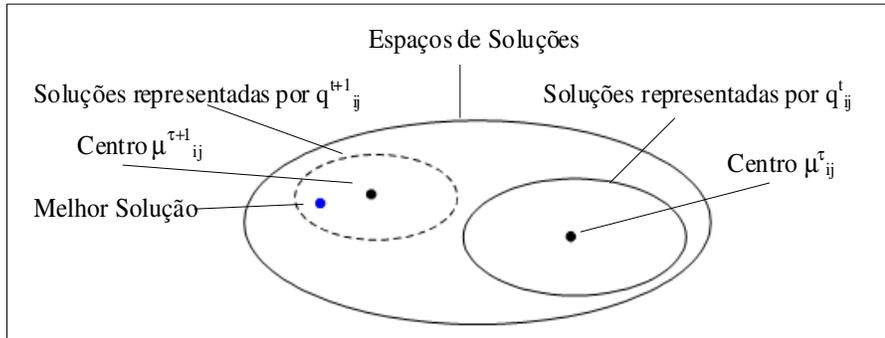


Figura 3.8 – Representação do espaço de soluções, e representação do espaço para um indivíduo q_{ij}^t e q_{ij}^{t+1} .

Sendo assim, esta seção objetivou esclarecer aspectos de atualização de indivíduos quânticos em $Q(t)$ com base na população dos melhores indivíduos clássicos $B(t)$. Na próxima seção, se apresentará como é feito o cruzamento clássico de indivíduos de $B(t)$ com os observados $P(t)$ de $Q(t)$ em uma geração t qualquer, no passo vi do algoritmo AEIQ-BR, conforme figura 3.1.

3.5. Cruzamento de Indivíduos Clássicos

A população em $B(t)$ funciona como um espaço de boas características, sendo refinada a cada passo t , através da introdução de novas soluções ajustadas de $B(t-1)$ e observadas de $P(t)$. Este refinamento de $B(t)$ acontece através da substituição de soluções no espaço de melhores soluções $B(t)$ por outras dominantes geradas através do cruzamento de $B(t)$ com $P(t)$, ou permanência em $B(t)$ caso o filho gerado não tenha um desempenho melhor. Este processo acontecerá através dos passos vi , vii e $viii$ do pseudocódigo em 3.1, e consistirá de:

- (1) Cruzamento clássico dos N indivíduos de $B(t-1)$ com os observados no passo atual em $P(t)$, gerando somente um filho para cada par de indivíduos. Cada par de indivíduos é selecionado aleatoriamente em $B(t-1)$ e $P(t)$ sem reposição;
- (2) Substituição dos N filhos gerados na população $P(t)$ atual;
- (3) Avaliação dos N filhos gerados em $P(t)$ anterior através de uma função de aptidão;
- (4) Atualização da população $B(t)$ com os N melhores indivíduos da união de duas populações, $B(t-1)$ e $P(t)$.

O mais importante nesta seção é entender como se dá este processo de cruzamento de indivíduos em $B(t)$ e $P(t)$. Este cruzamento será controlado por uma probabilidade de cruzamento clássico binário C_{Cb} para a parte binária, e outra C_{Cr} para a parte clássica real, ambas definidas pelo usuário e constante durante todas as gerações. Optou-se por desenvolver o algoritmo desta forma, pois se acredita na existência de dinâmicas de evolução distintas dadas por genes binários (combinatório) e reais (numérico) de um problema a ser otimizado, e, portanto, não podem ser controladas por uma única taxa.

O tipo de cruzamento a ser utilizado para a *parte binária* nos indivíduos clássicos será o *uniforme*, conforme se observa em Linden [32]. Esta decisão tem como objetivo minimizar efeitos de cruzamento de indivíduos, na presença de dependência ou correlação entre genes. Observar-se-á no seção 3.7 que este tipo de crossover será conveniente para evolução de redes neurais artificiais, e as decisões codificadas nos genes binários (alta dependência). Outro motivo pela escolha do cruzamento do tipo uniforme é análoga ao trabalho de Han e Kin [4] no algoritmo AEIQ-B.

O crossover para a parte real, por sua vez, foi o *aritmético* conforme o modelo AEIQ-R de Abs da Cruz [2]. Para facilitar a compreensão, suponham dois indivíduos clássicos originários das populações $B(t)$ e $P(t)$, respectivamente, representados por $[c^t_i]_B$ e $[c^t_i]_P$. Suponha também os genes $(x^t_{ij})_{bB}$ e $(x^t_{ij})_{bP}$, genes binários provenientes dos indivíduos $[c^t_i]_B$ e $[c^t_i]_P$. Pode-se definir um gene binário, $(x^t_{ij})_{bF}$, de um indivíduo filho $[c^t_i]_F$ da seguinte forma:

$$(x^t_{ij})_{bF} = \begin{cases} (x^t_{ij})_{bP}, & \text{se aleatório entre 0 e } C_{Cb} \\ (x^t_{ij})_{bB}, & \text{caso contrário} \end{cases}, \quad (3-21)$$

onde “*aleatório*” um número aleatório entre 0 e 1.

Analogamente para os genes reais $(x^t_{ij})_{rP}$ e $(x^t_{ij})_{rB}$ de $[c^t_i]_P$ e $[c^t_i]_B$, respectivamente, defini-se um gene real filho $(x^t_{ij})_{rF}$ de um indivíduo filho $[c^t_i]_F$ como a seguir:

$$(x^t_{ij})_{rF} = \begin{cases} (x^t_{ij})_{rB} + \lambda^t \left((x^t_{ij})_{rP} - (x^t_{ij})_{rB} \right), & \text{se aleatório entre 0 e } C_{Cr} \\ (x^t_{ij})_{rB}, & \text{caso contrário} \end{cases}, \quad (3-22)$$

onde “*aleatório*” um novo número aleatório entre 0 e 1, gerado para cada gene, e λ^t igual ao passo de atualização do indivíduo quântico na seção 3.4 anterior.

Um detalhamento mais técnico do processo (3-21) e (3-22) no passo vi do algoritmo AEIQ-BR pode ser observado pelo pseudocódigo na figura 3.9.

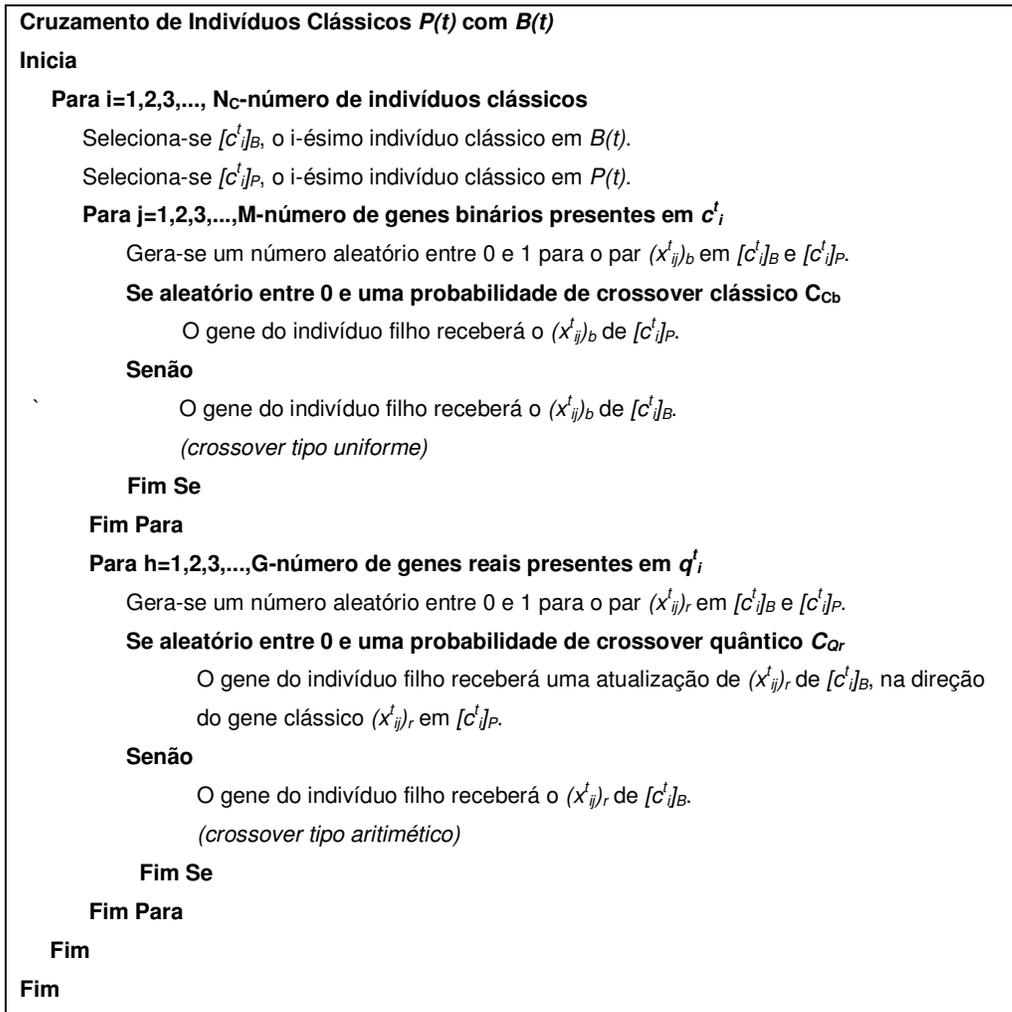


Figura 3.9 – Pseudocódigo do cruzamento de indivíduos clássicos em $P(t)$ com $B(t)$.

Para fins de exemplificação, suponham dois indivíduos clássicos selecionados para crossover conforme dado pela tabela 3.2. Suponha também probabilidades de crossover clássico binário $C_{Cb}=0,8$, e real $C_{Cr}=0,1$, e $\lambda^t=0,4$. Ou seja, a ativação e desativação de neurônios ocorreria em um ambiente mais dinâmico com alta probabilidade dada por C_{Cb} , enquanto que os pesos sofreriam ajustes mais finos, “sutis” com taxa da cruzamento dada por C_{Cr} , e λ^t uma velocidade de atualização considerável para os genes quânticos.

Tem-se que o novo indivíduo clássico a ser substituído em $P(t)$ no lugar de $[c_{i1}^t]_P$, o filho $[c_{i1}^t]_F$, pode ser definido como na figura 3.4.

$[c^t]_B$	$[c^t]_P$	Aleatório	$[c^t]_F$
$(x_{11})^t_b = 1$	$(x_{11})^t_r = 1$	0,627	$(x_{11})^t_r = 1$
$(x_{12})^t_b = 0$	$(x_{12})^t_b = 1$	0,718	$(x_{12})^t_b = 1$
$(x_{13})^t_b = 1$	$(x_{13})^t_b = 0$	0,758	$(x_{13})^t_b = 0$
$(x_{14})^t_b = 0$	$(x_{14})^t_b = 0$	0,726	$(x_{14})^t_b = 0$
$(x_{11})^t_r = 0,746$	$(x_{11})^t_r = 1,140$	0,363	$(x_{11})^t_r = 0,904$
$(x_{12})^t_r = 0,947$	$(x_{12})^t_r = -1,082$	0,992	$(x_{12})^t_r = 0,135$
$(x_{13})^t_r = 3,233$	$(x_{13})^t_r = 2,928$	0,092	$(x_{13})^t_r = 3,111$
$(x_{14})^t_r = -2,028$	$(x_{14})^t_r = 0,995$	0,684	$(x_{14})^t_r = -0,819$

Tabela 3.4 – Geração de um filho clássico.

A seguir, um resumo sobre o todo o detalhamento do algoritmo AEIQ-BR até o momento é oferecido, incluso ainda uma representação gráfica.

3.6. Esquema Gráfico e Resumo

Ao longo destas últimas seções, conclui-se que o algoritmo AEIQ-BR trata-se de um algoritmo evolutivo com alguns parâmetros ajustáveis pelo usuário. São eles:

- N_Q , número de indivíduos quânticos;
- N_C , número de indivíduos clássicos;
- C_Q , probabilidade de crossover quântico;
- C_{Cb} , probabilidade de crossover clássico binário;
- C_{Cr} , probabilidade de crossover clássico real;
- T , número de gerações ou evoluções;
- E θ , ângulo de rotação para velocidade de atualização dos genes clássicos binários.

Um quadro resumo dos passos do algoritmo AEIQ-BR pode ser visualizado conforme figura 3.10. Descritivamente tem-se o seguinte:

- (1) Em um passo inicial $t=1$, inicializa a população quântica $Q(t)$;
- (2) Gera-se uma população clássica a partir da observação de $P(t)$;
- (3) Avalia-se esta população clássica $P(t)$;
- (4) Se t for igual a 1, uma população de melhores $B(t)$ será gerada com $P(t)$;

- (5) Se t for maior do que 1, novos filhos serão substituídos em $P(t)$ com pais indivíduos clássicos gerados em t , $P(t)$, e melhores indivíduos obtidos em $t-1$, $B(t-1)$.
- (6) Avaliam-se estes filhos gerados armazenados em $P(t)$.
- (7) Melhores indivíduos da união de duas populações clássicas, $P(t)$ e $B(t-1)$, são armazenadas em $B(t)$.
- (8) Indivíduos quânticos $Q(t)$ são atualizados com melhores indivíduos de $B(t)$.
- (9) Enquanto t não atingir um máximo igual a T gerações, retorne ao passo 2) para que uma nova população clássica seja gerada.

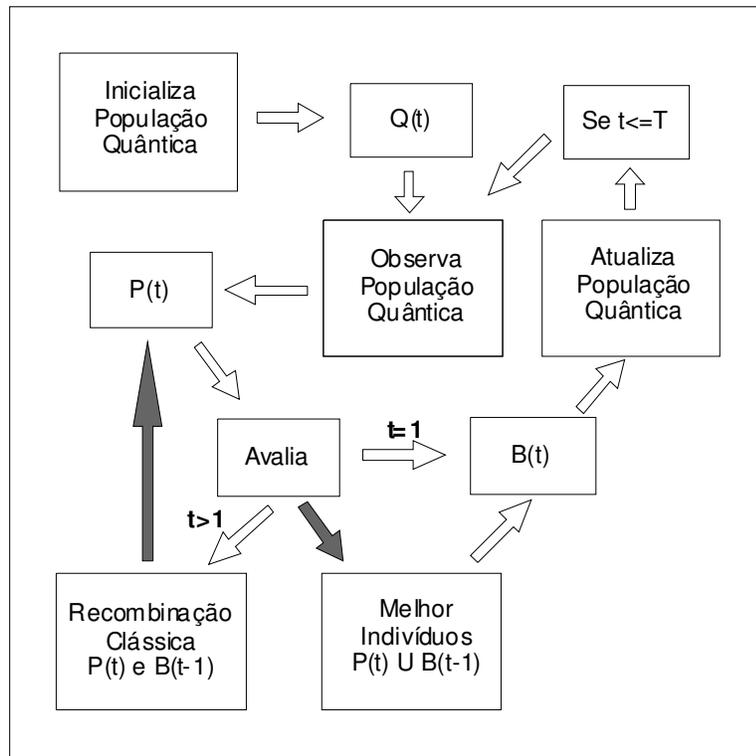


Figura 3.10 – Resumo do algoritmo AEIQ-BR.

Seja ainda uma população quântica com N indivíduos, cada indivíduo com M genes binários e G genes reais, e outra população clássica correspondente com H indivíduos, tem-se que estas populações podem ser representadas pela figura 3.11. A partir das próximas seções será importante que o leitor se sinta confiante com conceitos implícitos as figuras 3.10 e 3.11.

Na próxima seção em 3.7, se discutirá o algoritmo NEIQ-BR obtido da aplicação do algoritmo AEIQ-BR, para evolução de RNA do tipo alimentação adiante.

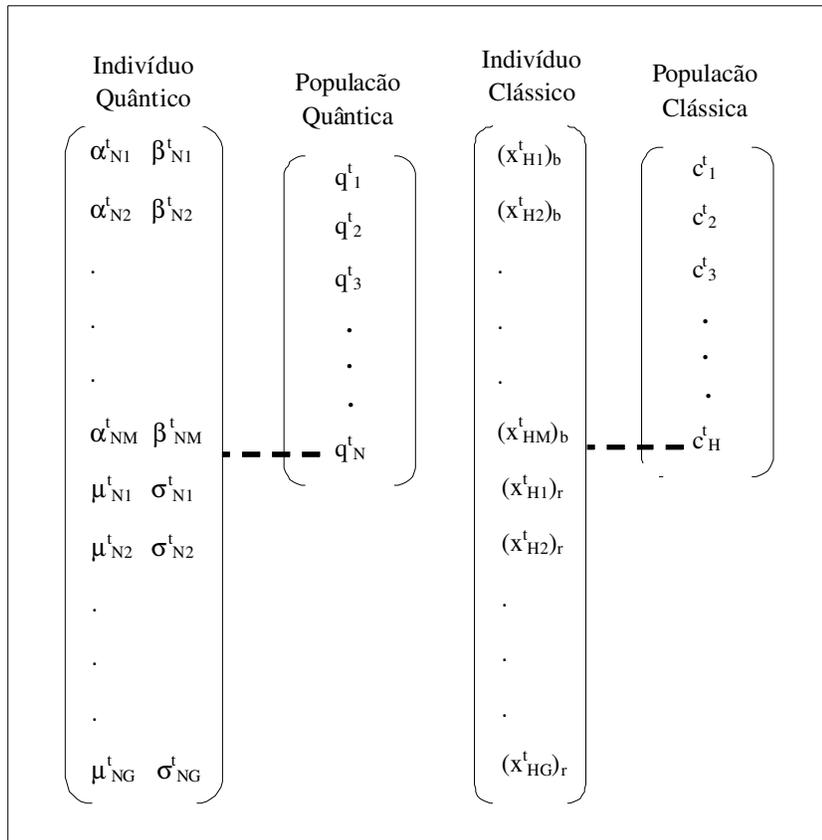


Figura 3.11 – Resumo da representação das populações Q(t) e P(t).

A seguir pode-se obter um quadro resumo para as principais características do algoritmo apresentado neste capítulo AEIQ-BR, em comparação com os algoritmos AEIQ-B e AEIQ-R.

Algoritmos	AEIQ-B	AEIQ-R	AEIQ-BR
Problemas	Combinatoriais	Numéricos	Ambos
Cromossoma			
Representação	Binária	Real	Ambos
Forma	Q-bit	Q-real	Ambos
Fundamentação	Notação Dirac	Pulso Quadrático	Ambos
Crossover			
Clássico Binário	Uniforme	-	Idem
Clássico Real	-	Aritmético	Idem
Quântico Binário	Porta Rotação	-	Idem
Quântico Real – Centro	-	Deslocamento Centro	Idem
Quântica Real – Amplitude	-	Regra do 1/5	Decaimento Exponencial

Tabela 3.5 – Principais características dos algoritmos AEIQ-B, AEIQ-R e AEIQ-BR.

3.7. Algoritmo NEIQ-BR

O principal objetivo em se utilizar o algoritmo AEIQ-BR nesta dissertação trata-se de configurar completamente uma rede neural do tipo perceptron com múltiplas camadas. Esta utilidade em se aplicar AEIQ-BR para evoluir uma RNA recebe o nome de algoritmo NEIQ-BR, numa extensão ao modelo de Abs da Cruz [2], o NEIQ-R, cuja representação da solução considerava somente características numéricas. A sigla NEIQ-BR vem de “**N**euro**E**volução **I**spirada na física **Q**uântica com representação mista **B**inária e **R**real”

Conforme mencionado no capítulo 1, esta modelagem *completa* de uma RNA pelo algoritmo NEIQ-BR, tendo em vista algum problema de classificação específico, será feita para responder os seguintes questionamentos:

- Quais variáveis utilizar na camada de entrada?
- Quantos neurônios utilizar na camada intermediária?
- Quais serão os pesos da camada de entrada?
- Quais serão os pesos da camada de intermediária?
- Qual função de ativação utilizar em cada neurônio da camada intermediária?
- Qual função de ativação utilizar em cada neurônio da camada de saída?

Decisões do tipo combinatórias e numéricas deverão ser tomadas e representadas, seja num indivíduo quântico, seja num indivíduo clássico. São decisões do tipo *combinatórias*: quais variáveis selecionar na camada de entrada; número de neurônios utilizados na camada escondida; e tipos de função de ativação a utilizar na camada escondida e saída. Por outro lado, são decisões do tipo *numéricas*: pesos da camada de entrada; e pesos da camada intermediária. Sobre estas decisões numéricas a serem otimizadas, *para qualquer variável de decisão x se considerará FDPs do tipo uniforme com domínio no intervalo de -4 a 4 .*

Nada mudará nas formas e métodos desenhados nas seções anteriores, quando detalhado o algoritmo AEIQ-BR. A generalização do NEIQ-BR se encontra principalmente na finalidade de aplicação do AEIQ-BR quando desejado agora evoluir uma RNA com os objetivos definidos anteriormente, mas também por algumas características a serem definidas na próxima seção.

Observaram-se no capítulo 2 que modelos neuroevolutivos são utilizados para problemas de diversas naturezas tais como: classificação, previsão, e controle. Neste trabalho, *a utilização do modelo NEIQ-BR terá como objetivo exclusivo o tratamento de problemas de classificação*, sendo a investigação no contexto de outros tipos de problemas envolvendo RNAs deixado para estudos futuros, como por exemplo, previsão de séries, controle e clusterização.

Conforme estudo apresentado na seção 2.2, tabela 2.5, não se encontrou outros autores que tenham desenvolvido um modelo NE capaz de lidar com todos os questionamentos levantados aqui, quando da modelagem de uma RNA para a classificação de padrões.

Algumas características do algoritmo NEIQ-BR são análogas ao algoritmo AEIQ-BR tais como: representação probabilística do espaço de busca; alta capacidade de exploração e aproveitamento; um otimizador global eficiente do problema; e, sobretudo, a capacidade de representação mista binário-real dos indivíduos.

3.7.1. Perceptrons de Múltiplas Camadas

O principal objetivo desta seção será compreender como é feito o processo de codificação e decodificação de um problema de modelagem de rede neural através de um cromossoma, seja ele quântico ou clássico, parte combinatória ou numérica. Primeiramente, as características da rede neural a ser modelada serão discutidas. Para isto, considere o seguinte:

- Seja um banco de dados com um total de **n observações**.
- Seja **TE**, o **total de variáveis independentes**.
- Seja **uma variável dependente** com um total de **TO classes**.
- Seja, **TO**, o **número de neurônios na camada de saída**, cada um representando uma classe da variável dependente.
- Seja **TI**, o **número de neurônios na camada intermediária**, um máximo permitido pelo usuário do algoritmo NEIQ-BR.
- Sejam as funções de ativação sigmóide **logística e hiperbólica**.

O tipo de rede neural a ser construída para qualquer problema de classificação pode ter representação dada pela figura 3.12.

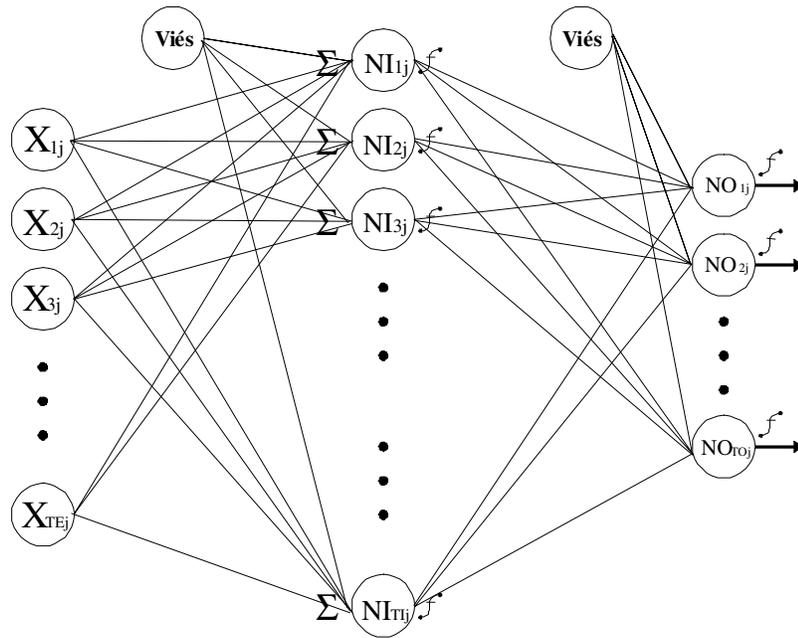


Figura 3.12 – Rede neural do tipo alimentação à frente.

Por esta figura, sinais são processados na camada de entrada para cada uma das n observações, $j=1,2,3,\dots,n$, e em cada um dos TI neurônios da camada intermediária, $i=1,2,3,\dots, TI$, ocorrendo uma soma ponderada conforme a seguir:

$$NI_{ij} = viés_i + \sum_{h=1}^{TE} X_{hj} \cdot w_{hi} . \quad (3-23)$$

E a ativação dos neurônios nesta camada intermediária da seguinte forma:

$$\varphi(NI_{ij}) = \begin{cases} \frac{1}{1 + \exp(-NI_{ij})}, & \text{se ativação logística.} \\ \frac{e^{NI_{ij}} - e^{-NI_{ij}}}{e^{NI_{ij}} + e^{-NI_{ij}}}, & \text{se ativação tangente hiperbólica.} \end{cases} . \quad (3-24)$$

A principal razão de escolher-se entre estas duas funções de ativação é que de acordo com Haykin [49], página 40, estas funções são as mais comumente utilizadas na modelagem de RNA. Analogamente, a ativação para cada neurônio TO na camada de saída pode ser obtida da seguinte forma:

$$NO_{ij} = viés_i + \sum_{h=1}^{TI} NI_{hj} \cdot w_{hi} , \quad (3-25)$$

onde $i=1,2,3,\dots,TO$, total de neurônios na camada escondida e $j=1,2,3,\dots,n$, total de observações. Vem-se que a ativação de cada neurônio na camada de saída pode ser obtida por:

$$\varphi(NO_{ij}) = \begin{cases} \frac{1}{1 + \exp(-NO_{ij})}, & \text{se ativação logística.} \\ \frac{e^{NO_{ij}} - e^{-NO_{ij}}}{e^{NO_{ij}} + e^{-NO_{ij}}}, & \text{se ativação tangente hiperbólica.} \end{cases} \quad (3-26)$$

Tendo em vista os resultados anteriores, segue-se com a representação de um indivíduo quântico para o caso de modelar uma RNA. Considerem um indivíduo quântico q^t_i , e um clássico c^t_h quaisquer. Para qualquer que seja o problema de classificação a ser modelado por uma RNA, a representação do indivíduo dependerá de:

- TE , o total de variáveis independentes
- TI máximo definido pelo usuário do algoritmo NEIQ-BR.
- TO , o número de neurônios na camada de saída

Sendo assim, pode-se definir q^t_i da seguinte forma:

$$q^t_i = \left[\left(q^t_i \right)_b \left(q^t_i \right)_r \right], \quad (3-27)$$

onde

$$\left(q^t_i \right)_b = \left(\alpha^t_{i1}; \dots; \alpha^t_{iTE}; \alpha^t_{i(TE+1)}; \dots; \alpha^t_{i(TE+TI)}; \alpha^t_{i(TE+TI+1)}; \dots; \alpha^t_{i(TE+2, TI)}; \alpha^t_{i(TE+2, TI+1)}; \dots; \alpha^t_{i(TE+2, TI+TO)} \right)_b \\ \left(\beta^t_{i1}; \dots; \beta^t_{iTE}; \beta^t_{i(TE+1)}; \dots; \beta^t_{i(TE+TI)}; \beta^t_{i(TE+TI+1)}; \dots; \beta^t_{i(TE+2, TI)}; \beta^t_{i(TE+2, TI+1)}; \dots; \beta^t_{i(TE+2, TI+TO)} \right)_b \quad (3-28)$$

e

$$\left(q^t_i \right)_r = \left(\mu^t_{i1}; \dots; \mu^t_{i(TE+(TI+1))}; \mu^t_{i(TE+(TI+1)+1)}; \dots; \mu^t_{i(TE+(TI+1)+(TI+1)*TO)} \right)_r \\ \left(\sigma^t_{i1}; \dots; \sigma^t_{i(TE+(TI+1))}; \sigma^t_{i(TE+(TI+1)+1)}; \dots; \sigma^t_{i(TE+(TI+1)+(TI+1)*TO)} \right)_r \quad (3-29)$$

Já c^t_h como:

$$c^t_h = \left[\left(c^t_h \right)_b \left(c^t_h \right)_r \right], \quad (3-30)$$

onde

$$\left(c^t_h \right)_b = \left(x^t_{h1}; \dots; x^t_{hTE}; x^t_{h(TE+1)}; \dots; x^t_{h(TE+TI)}; x^t_{h(TE+TI+1)}; \dots; x^t_{h(TE+2, TI)}; x^t_{h(TE+2, TI+1)}; \dots; x^t_{h(TE+2, TI+TO)} \right)_b \quad (3-31)$$

e

$$(c_h^t)_r = \left(x_{h1}^t; \dots; x_{h(TE \cdot (TI+1))}^t; x_{h(TE \cdot (TI+1)+1)}^t; \dots; x_{h(TE \cdot (TI+1)+(TI+1) \cdot TO)}^t \right)_r \quad (3-32)$$

De modo que $i=1, \dots, N$, número de indivíduos quânticos, e $h=1, \dots, H$ número de indivíduos clássicos. Com todos os indivíduos sujeitos as mesmas propriedades e manipulações do algoritmo AEIQ-BR detalhadas de 3.1 a 3.7. Desta forma, um indivíduo quântico ou binário teria um total de $TE+2 \cdot TI+TO$ genes binários, e $TE \cdot (TI+1)+(TI+1) \cdot TO$ genes reais.

A seguir uma figura resumo para a informação que é codificada em cada um dos genes quânticos ou clássicos, binários ou reais.

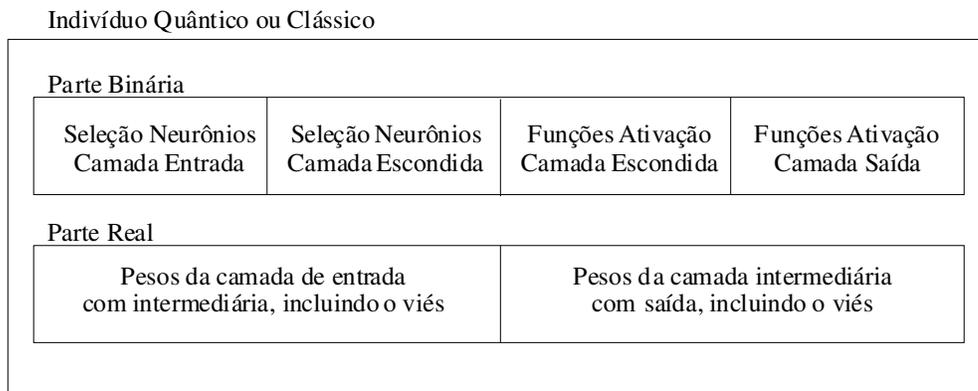


Figura 3.13 – Resumo de informação codificada para o modelo NEIQ-BR.

A decodificação de um cromossoma clássico pode ser resumida conforme a seguir:

- **Na camada de entrada**, códigos na parte binária do cromossoma determinarão quais variáveis serão selecionadas, ou seja, estarão ativas na RNA, onde o código 1 significará presença, e o código 0 significará ausência. Sendo assim, valores de entrada serão propagados adiante através de um subconjunto de um total de TE variáveis independentes, utilizando os pesos otimizados na parte real do cromossoma, ligação da entrada com a intermediária. Inclui-se também nestes pesos, o viés.
- **Na camada intermediária**, códigos na parte binária do cromossoma determinarão quais neurônios estarão ativos do total de TI , onde o código 1 significa a presença e código 0 significa a ausência. Cada neurônio ativo realizará a soma ponderada conforme (3-23) dos valores propagados da camada anterior. Após, cada um destes neurônios ativará esta soma ponderada com

função de ativação definida pela parte binária do cromossoma, onde 1 significará sigmóide logística e 0 tangente hiperbólica, conforme (3-24). Valores da camada intermediária serão propagados a diante com pesos codificados na parte real do cromossoma, ligação entra a camada intermediária e de saída. A parte real referente ao viés também estará sendo propagada.

- **Na camada de saída**, todos os neurônios estarão ativos, pois cada um representará uma de um total de TO classes. Todos os neurônios realizarão a soma ponderada dos valores propagados da camada anterior, e genes na parte binária do cromossoma indicarão a função de ativação a ser utilizada onde, da mesma forma, o código 1 significará sigmóide logística e 0 tangente hiperbólica.

Um detalhe importante é que o viés trata-se de um efeito sempre presente na RNA, e no cromossoma estará representado somente por genes na parte real, onde com a decodificação obtém pesos para as camadas intermediárias e de saídas.

A predição acontecerá através da atribuição da classe ao padrão de entrada representada pelo neurônio da camada de saída cuja ativação é máxima entre todos os TO . Esta associação para cada uma das n observações (ou padrões) será utilizada para o cálculo das funções de avaliação, também chamada de funções de aptidão do cromossoma.

3.7.2. Parâmetros de Configuração

No início da seção 3.6, apresentamos um resumo de todos os parâmetros de configuração do algoritmo AEIQ-BR até então. São eles: N_Q ; N_C ; C_Q ; C_{Cb} ; C_{Cr} , T ; e θ .

Em um indivíduo, conforme figura 3.13, os genes binários e reais apresentarão uma forte dependência entre si. A evolução e otimização dos pesos de um neurônio na parte real deve estar associada ao fato deste estar ativo ou não pela parte binária.

Seja um indivíduo clássico c^t_i qualquer e cruzamento clássico com taxa C_{Cr} , na ocasião de um neurônio estiver inativo pela decodificação de seu gene, é conveniente que não haja otimização de seus pesos correspondentes na parte real, pois a decodificação dos pesos em nada poderia se utilizar desta nova atualização.

Seja um indivíduo quântico q_i^t qualquer e cruzamento quântico com taxa C_Q , o problema de dependência entre genes com representações diferentes (real-binária) também estaria presente. O gene real x_{ij}^t – que codifica o peso de um neurônio qualquer – de um indivíduo clássico c_i^t não poderá ser utilizado para atualizar o gene correspondente na parte real do indivíduo quântico q_i^t , se o gene da parte binária do indivíduo clássico não estiver indicando uma atividade do neurônio que pertence x_{ij}^t .

Estes problemas tanto para o indivíduo clássico com crossover clássico e quântico com crossover quântico aconteceriam tanto na *camada de entrada*, quanto na *camada intermediária*.

Da mesma forma acontecerá para os tipos de funções de ativação aceitáveis por um neurônio qualquer, seja na camada intermediária, seja na camada de saída. Quando um neurônio estiver inativo não poderá ocorrer atualização com C_Q na parte binária quântica determinante dos tipos de funções de ativação de um indivíduo qualquer. E da mesma forma, na parte binária clássica de um indivíduo qualquer com taxa de cruzamento C_{Cb} , determinante dos tipos de funções.

Para resolver este problema, *o algoritmo AEIQ-BR foi modificado para que a evolução de pesos e funções de ativação nos cromossomos quânticos e clássicos esteja condicionada ao fato deste neurônio estar ativo ou não na parte binária correspondente*. Ou seja, os genes representantes dos pesos e funções de ativação permanecerão inalteráveis por cruzamento quânticos e clássicos, caso este neurônio esteja inativo. Como os cruzamentos clássicos utilizados são todos do tipo uniforme agindo com uma probabilidade independente entre genes, esta alteração em nada afeta a evolução de outros genes cujos neurônios estejam ativos⁵.

Feito esta última alteração de *condicionamento*, um alerta é verificado para que se possa precaver de possíveis conseqüências negativas no desempenho do NEIQ-BR.

Observem que uma vez que um neurônio esteja ativo pela decodificação de um gene da parte binária clássica de um indivíduo clássico, vindo de um longo período de inatividade, é necessário que se ocorra *um número consistente de atua-*

⁵ Observe que para um neurônio inativo que venha a ser ativado, é melhor que este continue a evoluir os seus pesos, pois não temos condições de saber se este neurônio sofreu um número suficiente de atualizações antes de ser inativado. E sendo assim, também não podemos garantir que reinicializá-lo seria uma estratégia eficiente. Somente um estudo empírico mais aprofundado poderia nos guiar, e, sobre isto, deixaremos para estudos futuros.

lização em seus pesos e função de ativação correspondente, para que através de um processo adaptativo este neurônio se torne relevante na rede neural.

O que se deseja que aconteça é: se este neurônio não se mostrar relevante para discriminar classes de um problema, que este assim se apresente após um longo processo de atualização de pesos e função de ativação. Não pode-se permitir que um neurônio não seja selecionado através do algoritmo evolutivo por este não ter sofrido atualizações suficiente em seus pesos sinápticos, quando de sua inatividade em boa parte do processo evolutivo.

Por estes motivos, um novo parâmetro passa a ser considerado para que o usuário do algoritmo possa ter um controle sobre este número consistente de gerações seguidas de atualização da parte real dos indivíduos quânticos e clássicos. Sendo assim, um primeiro novo parâmetro pode ser definido da seguinte forma:

- *mim* - o **número mínimo de gerações** que um neurônio permanecerá ativo antes que se **permita** a desativação por um cruzamento clássico, o que não significa que este assim será.

Este parâmetro fará com que haja um contador individual para cada gene clássico binário responsável pela atividade, ou inatividade, de um neurônio na camada de entrada e na camada intermediária. Conforme mencionado, após este contador – individual por indivíduo e genes de seleção de neurônios – atingir um número *mim* de gerações, estes genes passarão a ser considerados aptos para cruzamento clássicos.

Adicionalmente, com o objetivo de se ter algum controle entre aspectos de exploração e aproveitamento do espaço de soluções, um segundo novo parâmetro é considerado:

- *updateT* - o intervalo entre gerações até que os operadores quânticos sejam executado em todos os indivíduos de $Q(t)$, atualizando centros e amplitudes com melhores indivíduos de $B(t)$.

Ou seja, se deseja-se um processo gradual e lento de evolução para uma grande exploração do espaço de busca, este parâmetro deverá ser o mais alto. Enquanto que se deseja-se o aproveitamento de algum espaço de soluções, através de

um processo rápido de evolução, podendo incumbir em uma convergência prematura, este parâmetro deveria ser o mais baixo possível.

E por 3.7.1, como TE e TO serão definidos automaticamente pelo problema de classificação que se deseja modelar com a RNA, o usuário do NEIQ-BR também deverá considerar um terceiro novo parâmetro:

- TI , número máximo de neurônios permitidos na camada intermediária.

É importante dizer que todos os parâmetros devem ser ajustados de acordo com cada problema específico. Sendo assim, o usuário deverá estar atento a:

Parâmetros
N_Q
N_C
C_Q
C_{Cb}
C_{Cr}
T
θ
$\min T$
$\text{update } T$
TI

Tabela 3.6 – Parâmetros de ajuste do algoritmo NEIQ-BR.

A seguir, serão apresentadas algumas funções de aptidão a serem utilizadas juntamente com o algoritmo NEIQ-BR, e o desejo de se modelar problemas de classificação.

3.7.3. Funções de Avaliação ou Aptidão

Nesta seção, se apresentará um total de 3 formas de se avaliar um cromossoma clássico c^t_h no modelo NEIQ-BR, conforme passos *iv* e *vii* do pseudocódigo do AEIQ-BR, figura 3.1. Experimentos a serem definidos na seção 4.2 considerarão quando apropriado alguma destas 3 funções. São elas:

- (1) Percentual total de acerto ponderado, ou acurasse ponderada;
- (2) Percentual total de acerto não ponderado, ou acurácia;
- (3) E área abaixo da curva ROC.

A avaliação de um indivíduo c^t_h , pertencente a uma população $P(t)$, requer a leitura de toda uma base de dados e avaliação em termos de potencial de classificação. Para problemas de classificação dicotômica, a forma mais usual de se efe-

tuar esta avaliação é representá-la através da matriz de confusão. Suponha um problema qualquer com duas classes a serem preditas com um total de C_1 registros para a classe 1, e C_2 registros para a classe 2. Vem-se que a matriz de confusão para este problema dicotômico pode ser obtida da seguinte forma:

Classes		Classificada		
		C_1	C_2	Total
Observada	C_1	A_h^t	B_h^t	$A_h^t+B_h^t$
	C_2	C_h^t	D_h^t	$C_h^t+D_h^t$
	Total	$A_h^t+C_h^t$	$B_h^t+D_h^t$	$A_h^t+B_h^t+C_h^t+D_h^t$

Tabela 3.7 – Matriz de confusão para um problema de classes dicotômicas.

Onde A_h^t e D_h^t são os totais de registros classificados corretamente, respectivamente, *verdadeiro positivo* e *verdadeiro falso*. E C_h^t e B_h^t os erros de classificação, respectivamente, *falso positivo* e *falso negativo*.

Seja a razão C_1 dividida por C_2 denotado pela sigla $r_{c1/c2}$, ou seja, quantas vezes a classe C_1 é maior que a classe C_2 , pode-se obter a seguinte função de avaliação para um cromossoma c_h^t qualquer:

$$f(c_h^t) = \frac{(A_h^t + D_h^t \cdot r_{c1/c2})}{(A_h^t + B_h^t + C_h^t \cdot r_{c1/c2} + D_h^t \cdot r_{c1/c2})}, \tag{3-33}$$

ou seja, *o percentual total de acerto ponderado pela classe 1* através da razão $r_{c1/c2}$. Isto garantirá que o algoritmo evolutivo, mesmo na presença de números de elementos diferentes entre classes em um problema, busque soluções de RNAs que classificassem parcimoniosamente ambas as classes.

A generalização para a matriz de confusão na tabela 3.7, para um problema com n classes a serem preditas, pode ser obtido conforme a seguir:

Classes		Classificada					
		C_1	C_2	C_n	Total
Observada	C_1	a_{11}	a_{12}	a_{1n}	Σ
	C_2	a_{21}	a_{22}	a_{2n}	Σ
	Σ
	Σ
	C_n	a_{n1}	a_{n2}	a_{nn}	Σ
	Total	Σ	Σ	Σ	Σ	Σ	$\Sigma \Sigma$

Tabela 3.8 – Matriz de confusão para um problema de n classes.

Os índices t , geração corrente, e h , indivíduo clássico, vistos na tabela 3.7 foram omitidos na tabela 3.8 para simplificar a leitura. Seja i o índice de uma das

linhas da tabela 3.8, e j , das colunas, a generalização da função de avaliação dada em (3-33) para problemas com duas ou mais classes pode ser obtida por:

$$f(c_h^t)_{\% \text{ acerto total ponderado}} = \frac{\left(\sum_{i=j=1}^n a_{ij} \cdot r_{c1/ci} \right)}{\left(\sum_{i=1}^n \sum_{j=1}^n a_{ij} \cdot r_{c1/ci} \right)} \quad (3-34)$$

Adicionalmente, adota-se também a forma mais simples de (3-34), a função *percentual de acerto total não ponderado*, a qual pode ser obtida simplesmente excluindo-se o termo de razão da equação acima, sendo também considerada em nossos estudos. Tem-se então que:

$$f(c_h^t)_{\% \text{ acerto total}} = \frac{\left(\sum_{i=j=1}^n a_{ij} \right)}{\left(\sum_{i=1}^n \sum_{j=1}^n a_{ij} \right)} \quad (3-35)$$

Com isto, é esperado que soluções que classifiquem corretamente a classe predominante em um problema qualquer sejam favorecidas em gerações iniciais por apresentarem uma função de avaliação mais elevada. Em outras palavras, é conveniente para um algoritmo que não considera custos por classes associados ao total de registros preditos erroneamente, apresentar uma leve tendência em se “especializar” em classes predominantes.

Um exemplo disto é o algoritmo de retropropagação com método do gradiente decrescente. Um usuário interessado em otimizar os pesos de uma rede através deste algoritmo deve se preocupar em se ter uma amostra balanceada entre classes de interesse de previsão. Do contrário – caso os custos de classificação entre classes sejam iguais – o usuário se encontrará diante de sérios problemas de generalização e parcimônia entre classes, podendo a rede treinada se especializar em alguma das classes predominantes.

Contudo, considerou-se esta função de avaliação, o percentual de acerto não ponderado, e na oportunidade, resultados serão comparados pelo percentual de acerto ponderado quando obtidos.

E ainda, se considerará uma terceira função de avaliação, além das percentuais totais de acerto ponderado e não ponderado. Trata-se da *área abaixo da cur-*

va *ROC*⁶ – em inglês *Receiver Operating Characteristic* – conforme se observa em Braga [34]. Segundo Bradley, a área sob a curva ROC surge numa necessidade de se ter uma medida única de desempenho invariante com probabilidades a priori, e de fácil extensão para consideração análise de custo e benefício.

Uma curva ROC é obtida através de dois conceitos, o de sensibilidade, e o de especificidade. Sensibilidade é dita como a fração de padrões classificados como positivos que são verdadeiramente positivos, ou *FVP* e será representada através da coordenada de um plano cartesiano. E (1-especificidade) é dita como a fração de padrões classificados como falsos positivos, ou *FFP*, e será representada através da abscissa deste mesmo plano. Considerando a tabela 3.7, os valores respectivos para *FVP* e *FFP* podem ser obtidos como:

$$\begin{aligned} (\text{sensibilidade}) = FVP &= \frac{A_h^t}{A_h^t + C_h^t}, \\ (1 - \text{especificidade}) = FFP &= \frac{B_h^t}{B_h^t + D_h^t} \end{aligned} \quad (3-36)$$

Seja um indivíduo c_h^t qualquer, com somente um neurônio na camada de saída, a classificação de um padrão em uma de duas classes aconteceria através da atribuição de um corte ou limiar na pseudo-probabilidade deste neurônio. Em caso da ativação acima deste limiar, classifica-se este padrão como positivo, do contrário, negativo.

Não é necessário que o usuário atribua algum ponto de corte para qualquer pseudo-probabilidade, pois faz parte do cálculo da área da curva ROC a variação deste limiar um total de n vezes no intervalo de 0 até 1. Sendo assim, variando-se este ponto de corte n vezes de 0 até 1, n pares de *FVP* e *FFP* poderão ser obtidos, e serão representados como no gráfico 3.14.

A reta que passa pela diagonal representa o resultado de um modelo com base numa seleção aleatório para classificação de registros entre classes C_1 e C_2 . A aptidão de um indivíduo c_h^t , por sua vez, seria obtida pela área abaixo da curva ROC até a reta que passo pela diagonal. Ou seja, o ganho que se tem de sair de um modelo aleatório, para um probabilístico.

⁶ Em caso da função de avaliação ser a Curva ROC, somente um neurônio será considerado na camada de saída do cromossoma, e indivíduos quânticos e binários seriam representados em concordância com esta camada.

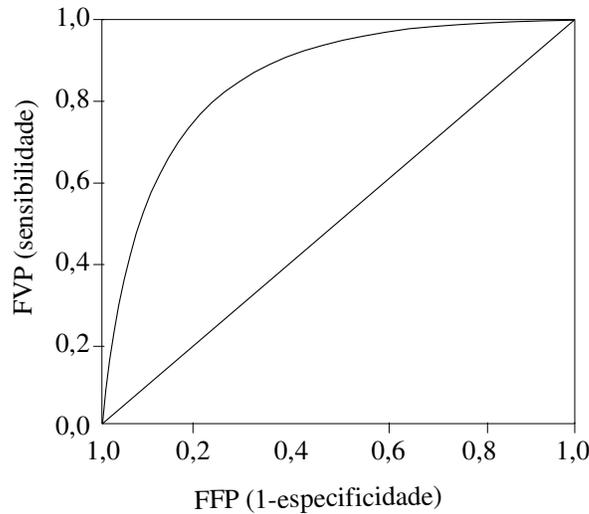


Figura 3.14 – Curva ROC.

Conforme se observa em Braga [34], esta área abaixo da curva ROC pode ser obtida de através de diversas formas como: regra do trapézio; estimação de máxima verossimilhança; aproximação pela estatística U de Wilcoxon-Mann-Whitney; e a partir do declive e termo de intercepção da representação dos dados originais em papel de probabilidades binormais. Nesta dissertação, adotar-se-á a *regra do trapézio* como cálculo da área sob a Curva ROC.

Sejam os conceitos de sensibilidade e especificidade escritos conforme (3-36). Quando o limiar (em inglês, *threshold*) é variado ao longo de seu domínio de 0 até 1, os pontos de (sensibilidade, especificidade) são registrados. Neste caso, uma maneira simples de se obter a área sobre a curva pela *regra do trapézio* seria como:

$$Área = \sum_{i=1}^n \left\{ (1 - (1 + FVP_i) \cdot \Delta FFP) + \frac{1}{2} [\Delta FVP \cdot \Delta FFP] \right\}, \quad (3-37)$$

onde

$$\begin{aligned} \Delta FVP &= FVP_i - FVP_{i-1}, \\ \Delta FFP &= FFP_i - FFP_{i-1}, \end{aligned} \quad (3-38)$$

e n , o número de variações do ponto de corte.

Uma curva ROC, por definição, se aplica somente a problemas de classificação binária, pois é fundada em princípios de falso positivos e negativos, e verdadeiros positivos e negativos. Por tais motivos, esta última função de avaliação não será utilizada em problemas com mais de duas classes, conforme casos de estudos a serem apresentados no capítulo 4.

Sendo assim, a função conforme (3-37) será adotada como uma medida de aptidão para problemas onde existem somente duas classes a serem preditas. Sendo assim, utilizar-se-á três formas de se avaliar um cromossoma clássico em termos de desempenho obtido em um problema de classificação qualquer, conforme discutido anteriormente. Tais funções serão utilizadas em estratégias evolutivas nos passos *iv* e *vii* do algoritmo NEIQ-BR, em experimentos a serem detalhados na seção 4.2. São elas:

- (1) Percentual total de acerto ponderado pela classe 1, conforme (3-34);
- (2) Percentual total de acerto não ponderado, conforme (3-35);
- (3) E área por debaixo da curva ROC, conforme (3-37).